4-1-1990

Tests for Parameter Instability and Structural Change with Unknown Change Point

Donald W.K. Andrews

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY

Box 2125, Yale University
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 943

TESTS FOR PARAMETER INSTABILITY AND STRUCTURAL CHANGE WITH UNKNOWN CHANGE POINT

by

Donald W.K. Andrews

October 1989

Revised: April 1990

TESTS FOR PARAMETER INSTABILITY AND STRUCTURAL CHANGE WITH UNKNOWN CHANGE POINT

Donald W. K. Andrews[1]

Cowles Foundation for Research in Economics
Yale University

October 1989

Revised: April 1990

ABSTRACT

This paper considers tests of parameter instability and structural change with unknown change point. The results apply to a wide class of parametric models including models that satisfy maximum likelihood–type regularity conditions and models that are suitable for estimation by generalized method of moments procedures. The paper considers likelihood ratio and likelihood ratio–like tests, as well as asymptotically equivalent Wald and Lagrange multiplier tests. Each test implicitly uses an estimate of a change point. Tests of both "pure" and "partial" structural change are discussed.

The asymptotic distributions of the test statistics considered here are non–standard because the change point parameter only appears under the alternative hypothesis and not under the null. The asymptotic null distributions are found to be given by the supremum of the square of a standardized tied–down Bessel process of order $p \geq 1$. Tables of critical values are provided based on this asymptotic null distribution.

As tests of parameter instability, the tests considered here are shown to have non-trivial asymptotic local power against all alternatives for which the parameters are non–constant. As tests of one–time structural change, the tests are shown to have some weak asymptotic local power optimality properties for large sample size and small significance level. The tests are found to perform quite well in a Monte Carlo experiment.

JEL Classification No.: 211

Keywords: Asymptotic distribution, change point, Brownian bridge, Brownian motion, F–test, generalized method of moments estimator, Lagrange multiplier test, likelihood ratio test, maximum likelihood estimator, parameter instability, structural change, Wald test, weak convergence.

1. INTRODUCTION

This paper considers tests of parameter instability and structural change with unknown change point. The proposed tests are designed for a one–time change in the value of a parameter vector, but are shown to have power against general forms of parameter instability. In particular, the paper considers the likelihood ratio (LR) or LR–like test for one-time structural change with unknown change point, as well as the analogous Wald (W) and Lagrange multiplier (LM) tests. The tests considered here apply quite generally to parametric models that do not exhibit deterministic or stochastic trends.[2]

The results of the paper apply to tests of "pure" structural change and "partial" structural change. With pure structural change, the entire parameter vector is allowed to exhibit parameter instability across the observations under the alternative hypothesis. With partial structural change, only a subvector of the parameter vector is subject to instability under the alternative hypothesis.

The statistical literature on change point problems is extensive. (See "change point" in *Current Index to Statistics* and the review papers by Zacks (1983) and Krishnaiah and Miao (1988). Recent references include James, James, and Siegmund (1987), D. L. Hawkins (1987), and Kim and Siegmund (1989) among others.) Many of the results in the literature concern location models, other scalar parameter models, or simple regression models. Most of the results only apply to tests of pure structural change. The most general results available appear to be those of D. L. Hawkins (1987), who considers Wald tests of pure structural change based on maximum likelihood (ML) estimators in independent identically distributed (iid) scenarios. These results are still too restrictive, however, for many econometric applications. In consequence, in this paper we establish results that allow for dependent non–identically distributed (dnid) observations, estimation by methods other than ML, tests of both pure and partial structural change, and tests based on W, LM, and LR (or LR–like) test statistics. For example, stationary models estimated by generalized method of moments (GMM) and dnid models estimated by ML are used to illustrate the general results.

We now introduce and motivate the tests considered in this paper. With the linear regression model, it is common in econometrics to test for a one–time structural change occurring at a given time point using an F–test (often referred to as a Chow test). This test can be extended to W, LM–like, and LR–like tests in general parametric models, whether estimation is by maximum likelihood or not, see Andrews and Fair (1988). For specificity, let $W_T(\pi)$, $LM_T(\pi)$, and $LR_T(\pi)$ denote the W, LM–like, and LR–like tests referred to above, where T is the sample size, $[T\pi]$ is the given change point for some $\pi \in (0,1)$, and $[T\pi]$ denotes the integer part of $T\pi$. For simplicity, we refer to π, rather than $[T\pi]$, as the point of structural change.

Although the tests $W_T(\pi), \ldots, LR_T(\pi)$ are widely used and exhibit various optimality properties, there are several commonly occurring situations in which these tests are not appropriate. It is the purpose of this paper to consider alternative, but closely related, tests for these situations. These situations include the following: First, suppose one has no prior information regarding the point π at which structural change may occur. Then, the tests $W_T(\pi), \ldots, LR_T(\pi)$ can be applied only by using some ad hoc choice of π. Such a choice has adverse effects on the power of the tests, since the change point is misspecified for many alternatives of interest.

Second, suppose a suitable value of π is available, but this value has been determined by looking at the data or by looking at data that is not independent of the data on which the test is to be applied. Then, the tests $W_T(\pi), \ldots, LR_T(\pi)$ are not valid for reasons of data–mining. That is, the sizes of the tests are not correct even in large samples. Note that for macroeconomic applications especially this problem is likely to be prevalent, since many different applications use the same or similar data and most macroeconomic variables are correlated with one another.

Third, suppose a suitable value of π is available and has been determined by an "exogenous" event, such as an oil price shock. Then, the tests $W_T(\pi), \ldots, LR_T(\pi)$ still are not valid for data–mining reasons if the exogenous event has been chosen endogenously. For example, suppose there are multiple exogenous events that occur during the sample period, any one of which has the potential to induce structural change. If one chooses a particular exogenous event on the basis of empirical studies that use the same or similar data or data that are not independent of those to be used for the test, then the resulting test is again subject to the criticism of data–mining.

Fourth, if one is interested in testing against more general forms of structural change than one–time changes, then a test that uses a fixed change point is likely to have low or no power against many alternatives of interest. Thus, as tests of parameter instability, the tests $W_T(\pi), \ldots, LR_T(\pi)$ are inadequate.

The obvious solution to the inadequacy of $W_T(\pi), \ldots, LR_T(\pi)$ in the situations described above is to treat the change point π as an unknown and to construct tests for structural change that do not take π as given. The problem of testing for one–time structural change with an unknown change point, however, does not fit into the standard "regular" testing framework, see Davies (1977, 1987). The reason is that the parameter π only appears under the alternative hypothesis and not under the null. In consequence, W, LM, and LR tests constructed with π treated as a parameter do not possess their standard large sample asymptotic distributions.

In this paper, we adopt a common method used in this scenario and consider test statistics of the form

$$\text{(1.1)} \qquad \sup_{\pi\in\Pi} W_T(\pi)\,, \ \sup_{\pi\in\Pi} LM_T(\pi)\,, \ \text{and} \ \sup_{\pi\in\Pi} LR_T(\pi)\,,$$

where Π is some pre–specified subset of [0,1] whose closure lies in (0,1).[3] (The reasons for taking Π as such are discussed below.) Tests of this form can be motivated or justified along several grounds. First, $\sup_{\pi\in\Pi} LR_T(\pi)$ is the LR (or LR–like) test statistic for the case of unspecified parameter π with parameter space Π. In addition, the test statistics

$\sup_{\pi\in\Pi} W_T(\pi)$ and $\sup_{\pi\in\Pi} LM_T(\pi)$ generally are asymptotically equivalent to $\sup_{\pi\in\Pi} LR_T(\pi)$ under the null and local alternatives. Second, the test statistics $\sup_{\pi\in\Pi} W_T(\pi)$, $\ldots$, $\sup_{\pi\in\Pi} LR_T(\pi)$ correspond to the tests derived from Roy's type I (or union-intersection) principle, see Roy (1953) and Roy, Gnanadesikan, and Srivastava (1971, pp. 36–46). Third, the above test statistics can be shown to possess certain (weak) asymptotic optimality properties against local alternatives for large sample size and small significance level. These results are due to Davies (1977, Thm. 4.2) for scalar parameter one–sided tests and are extended below to multi–parameter two–sided tests.

Below we determine the asymptotic distributions of the test statistics of (1.1) under the null hypothesis of parameter constancy and under local alternatives of parameter instability including one–time structural change. Specifically, the local alternatives considered are of the form $\theta_1 = \theta_{10} + \frac{1}{\sqrt{T}}\,\eta(t/T)$ for some bounded function $\eta(\cdot)$ on [0,1]. The local power results of the paper show that the tests of (1.1) have power against all local alternatives for which $\eta(\cdot)$ is not almost everywhere on Π equal to a constant. Thus, as tests of parameter instability, the tests of (1.1) have some desirable properties.

As tests of parameter instability, the tests of (1.1) can be compared with several other tests in the literature, such as the CUSUM and CUSUM of squares tests of Brown, Durbin, and Evans (1975) and the fluctuation test of Sen (1980) and Ploberger, Krämer, and Kontrus (1989). These tests are all designed for the linear regression model. A drawback of the CUSUM test is that it exhibits only trivial power against alternatives in certain directions, as shown by Krämer, Ploberger, and Alt (1988) using asymptotic local power and by Garbade (1977) and others using simulations. In addition, the CUSUM of squares test has only trivial asymptotic local power against all alternatives of the form discussed above, see Ploberger and Krämer (1986). In contrast, the tests of (1.1) do not exhibit these local power problems.

The fluctuation test is similar to the $\sup_{\pi\in\Pi} W_T(\pi)$ test of (1.1), but the latter possesses large sample optimality properties for each fixed π, whereas the former does not. The reason is that the fluctuation test aggregates the elements of a vector by taking the supremum over the elements, whereas the $\sup_{\pi\in\Pi} W_T(\pi)$ test aggregates the elements in the standard way using a quadratic form. In addition, the fluctuation test statistic $\sup_{\pi\in\Pi} FL_T(\pi)$ is unequally weighted across different values of π, since the asymptotic variance of $FL_T(\pi)$ is $\pi(1-\pi)$ under the null hypothesis. In contrast, the $\sup_{\pi\in\Pi} W_T(\pi)$ statistic is equally weighted across different values of π, since the asymptotic variance of $W_T(\pi)$ is one for all $\pi \in (0,1)$ under the null hypothesis. In sum, the differences outlined above suggest that the $\sup_{\pi\in\Pi} W_T(\pi)$ test may have better all around power properties than the fluctuation test.

We note that the CUSUM, CUSUM of squares, and fluctuation tests have been analyzed in the literature only in the context of the linear regression model. In contrast, the results given here for the tests of (1.1) apply to a general class of models. In addition, they apply to tests of both pure and partial structural change. The results given here also can be used to extend the applicability of the fluctuation test to more general models and to tests of partial structural change, for details see Andrews (1989e).

Several additional tests in the literature for testing for parameter instability are the tests of Leybourne and McCabe (1989) and Nyblom (1989). (Also, see the references in Krämer and Sonnberger (1986, pp. 56–59).) These tests are designed for non–stationary alternatives, and hence, have a different focus than the tests considered here.

The focus of this paper is on formal hypothesis tests. Nevertheless, the statistics, statistical processes, and asymptotic distribution theory discussed here can be used quite effectively in a data analytic fashion. In particular, they can be used in a manner similar to that advocated by Brown, Durbin, and Evans (1975) for recursive residuals and the CUSUM test and by Hendry (1989, pp. 44, 49) for rolling change point tests.

The remainder of this paper is organized as follows: Section 2 introduces a class of partial–sample extremum (PSE) estimators, establishes their consistency, and determines their asymptotic distributions. Section 3 introduces W, LM, and LR tests of parameter instability based on the PSE estimators. Section 4 determines the asymptotic null distributions of the W, LM, and LR test statistics and provides tables of critical values for them. Section 4 also establishes the asymptotic distributions of these test statistics under local alternatives and obtains two local power optimality results. In Sections 2–4, the GMM and ML cases are treated explicitly as examples of the general results. Section 5 presents some simulation evidence regarding the performance of the tests in the linear regression context. An Appendix provides proofs of the results given in the paper.

Lastly, we mention several notational conventions that are used throughout the paper: Unless specified otherwise, all limits are taken as $T \to \infty$, where T is the sample size. The symbol $\Rightarrow$ denotes weak convergence (as defined by Billingsley (1968, Ch. 1) using the Skorohod metric or by Pollard (1984, pp. 64–66) using the uniform metric), $\xrightarrow{d}$ denotes convergence in distribution, $\xrightarrow{p}$ denotes convergence in probability, Σ_a^b abbreviates $\Sigma_{t=a}^b$, $\|\cdot\|$ denotes the Euclidean norm, and for simplicity $T\pi$ denotes $[T\pi]$, where $[\cdot]$ is the integer part operator. Π denotes a set whose closure lies in (0,1).

2. PARTIAL–SAMPLE EXTREMUM ESTIMATORS

In this section we analyze partial–sample extremum (PSE) estimators. PSE estimators are extremum estimators that primarily use the pre–$T\pi$ or the post–$T\pi$ data in estimating a parameter θ_1 for variable values of π in Π and use all the data in estimating an additional parameter θ_3. These estimators are the basic components of the Wald test of parameter instability and structural change with unknown change point. Furthermore, the properties of PSE estimators are used to obtain the asymptotic distributions of the corresponding LM–like and LR–like tests.

The first subsection below defines the class of estimators to be considered. The second provides conditions under which they are consistent uniformly over $\pi \in \Pi$. The third subsection establishes the weak convergence of PSE estimators to a function of a vector Brownian motion process on [0,1] restricted to Π. The fourth subsection considers the estimation of unknown matrices that arise in the limiting Brownian motion process. GMM and ML estimators are discussed throughout to exemplify the results given.

2.1. Definition of Partial–Sample Extremum Estimators

The data are given by a triangular array of random vectors (rv's) $\{W_{Tt}\} = \{W_{Tt} : t = 1, \ldots, T; T \geq 1\}$ defined on a probability space $(\Omega, \mathcal{B}, P)$. The observed sample is $\{W_{Tt} : t = 1, \ldots, T\}$. Often we let W_t abbreviate W_{Tt}. PSE estimators are defined as follows:

DEFINITION: A sequence of *partial–sample extremum* (PSE) estimators $\{\hat{\theta}(\cdot) : T \geq 1\} = \{(\hat{\theta}(\pi) : \pi \in \Pi) : T \geq 1\}$ is any sequence of stochastic processes such that

$$d(\bar{m}_T(\hat{\theta}(\pi), \pi, \hat{\underset{\sim}{\tau}}(\pi)), \hat{\gamma}(\pi)) = \inf_{\theta \in \Theta} d(\bar{m}_T(\theta, \pi, \hat{\underset{\sim}{\tau}}(\pi)), \hat{\gamma}(\pi)) \tag{2.1}$$

for all $\pi \in \Pi$ with probability $\to 1$, where $\theta = (\theta_1', \theta_2', \theta_3')' \in \Theta_1 \times \Theta_1 \times \Theta_3 \subset R^p \times R^p \times R^{p_3}$,
$$\bar{m}_T(\theta, \pi, \underset{\sim}{\tau}) = \frac{1}{T} \Sigma_1^{T\pi} \begin{bmatrix} m_{1t}(\theta_1, \theta_3, \tau_1) \\ \underset{\sim}{0} \\ m_{3t}(\theta_1, \theta_3, \tau_1) \end{bmatrix} + \frac{1}{T} \Sigma_{T\pi+1}^{T} \begin{bmatrix} \underset{\sim}{0} \\ m_{1t}(\theta_2, \theta_3, \tau_2) \\ m_{3t}(\theta_2, \theta_3, \tau_2) \end{bmatrix} \in R^v,$$
$\underset{\sim}{\tau} = (\tau_1', \tau_2')' \in \mathcal{T} \times \mathcal{T} \subset R^u \times R^u$, $m_{rt}(\theta_1, \theta_3, \tau_1)$ abbreviates $m_{rTt}(W_{Tt}, \theta_1, \theta_3, \tau_1)$ for $r = 1, 3$, $m_{rTt}(\cdot,\cdot,\cdot,\cdot)$ is a function from $R^{k_{Tt}} \times \Theta_1 \times \Theta_3 \times \mathcal{T}$ to R^{v_r} for $r = 1, 3$, k_{Tt} is a positive integer $\leq \infty$, $\hat{\underset{\sim}{\tau}}(\pi) = (\hat{\tau}_1(\pi)', \hat{\tau}_2(\pi)')'$ and $\hat{\gamma}(\pi)$ are random elements of $\mathcal{T} \times \mathcal{T}$ and $\Gamma \subset R^g$ (which depend on T in general), and $d(\cdot,\cdot)$ is a non–random, real–valued function (which does not depend on T).

As the definition indicates, $\hat{\theta}(\pi) = (\hat{\theta}_1(\pi)', \hat{\theta}_2(\pi)', \hat{\theta}_3(\pi)')'$ is a $2p+p_3$–vector comprised of an estimator $\hat{\theta}_1(\pi) \in R^p$ that primarily uses the pre–$T\pi$ data, an estimator

$\hat{\theta}_2(\pi) \in R^p$ that primarily uses the post–$T\pi$ data, and an estimator $\hat{\theta}_3(\pi) \in R^{p_3}$ that uses all of the data. Under the null hypothesis of parameter constancy $\hat{\theta}_1(\pi)$ and $\hat{\theta}_2(\pi)$ are estimators of the same parameter, call it θ_{10}. In the case of tests of pure structural change, there is no parameter θ_3 or function $m_{3t}(\theta_1, \theta_3, \tau_1)$. In this case, $\hat{\theta}_1(\pi)$ is based strictly on the pre–$T\pi$ data and $\hat{\theta}_2(\pi)$ strictly on the post–$T\pi$ data. In the case of tests of partial structural change, the parameter θ_3 appears and is taken to be constant across the observations under both the null hypothesis and the alternative. The preliminary nuisance parameter estimators $\hat{\underset{\sim}{\tau}}(\pi)$ and $\hat{\gamma}(\pi)$ that appear in the criterion function are allowed to depend on π, but this dependence is often suppressed below for notational simplicity.

For a fixed value of π, the PSE estimators defined above are quite similar to the extremum estimators of Andrews and Fair (1988). For $\pi = 1$, the PSE estimators are analogous to the estimators analyzed in Andrews (1989a,b,c) but with the restriction that $\hat{\underset{\sim}{\tau}}(\pi)$ is finite dimensional. In fact, at the cost of greater complexity, the stochastic equicontinuity approach espoused in the latter papers can be employed here as well. Thus, it is possible to extend many of the results given below to semiparametric PSE estimators and to PSE estimators that are based on functions $m_{rt}(\theta_1, \theta_3, \tau_1)$, $r = 1, 3$, that are not differentiable in θ_1 or θ_3.

Next, we introduce two examples of PSE estimators:

EXAMPLE 1 (GMM): In this example, we discuss the generalized method of moments (GMM) estimator as defined and analyzed by Hansen (1982). We consider the case of pure structural change. That is, no parameter θ_3 or function $m_{3t}(\theta_1, \theta_3, \tau_1)$ appears in the definition of $\hat{\theta}(\pi)$. The pre–$T\pi$ GMM estimator $\hat{\theta}_1(\pi)$ is defined to minimize the quadratic form

$$\left[\frac{1}{T}\Sigma_1^{T\pi} f(W_t, \theta_1)\right]' \frac{1}{\pi} a_{1,T\pi}' a_{1,T\pi} \left[\frac{1}{T}\Sigma_1^{T\pi} f(W_t, \theta_1)\right] / 2 \tag{2.2}$$

over $\theta_1 \in \Theta_1 \in R^p$, where $a_{1,T\pi}' a_{1,T\pi}$ is a weight matrix defined using the pre–$T\pi$ data and $f(\cdot,\cdot)$ is a given function that satisfies the orthogonality conditions $Ef(W_t, \theta_{10}) = \underset{\sim}{0}$ $\forall t \geq 1$ for some $\theta_{10} \in \Theta_1$ when no parameter instability or structural change occurs. The post–$T\pi$ GMM estimator $\hat{\theta}_2(\pi)$ is defined analogously with $\Sigma_1^{T\pi}$, $1/\pi$, and $a_{1,T\pi}$ replaced by $\Sigma_{T\pi+1}^{T}$, $1/(1-\pi)$, and $a_{2,T\pi}$, respectively, where $a_{2,T\pi}' a_{2,T\pi}$ is a weight matrix defined using the post–$T\pi$ data.

The GMM estimator is easily seen to be a PSE estimator with

$$m_{1t}(\theta_1, \tau_1) = f(W_t, \theta_1)\ , \quad \hat{\gamma}(\pi) = \mathrm{diag}\{a_{1,T\pi}' a_{1,T\pi}/\pi,\ a_{2,T\pi}' a_{2,T\pi}/(1-\pi)\}\ ,$$
$$\text{and} \quad d(m,\gamma) = m'\gamma m/2\ . \tag{2.3}$$

In this example, no nuisance parameter estimator $\hat{\underset{\sim}{\tau}}(\pi)$ appears.

EXAMPLE 2 (ML): This example considers maximum likelihood (ML) estimators in models with endogenous and exogenous variables and with data that may be dnid. We consider partial structural change, so $\hat{\theta}(\pi)$ contains a sub–vector $\hat{\theta}_3(\pi)$. For $t = 1, \ldots, T$, let

$$\begin{aligned} &\{f_t(\theta_1, \theta_3) : \theta_1 \in \Theta_1,\ \theta_3 \in \Theta_3\} \\ &\qquad = \{f_t(Y_t | Y_1, \ldots, Y_{t-1}; X_1, \ldots, X_t; \theta_1, \theta_3) : \theta_1 \in \Theta_1,\ \theta_3 \in \Theta_3\} \end{aligned} \tag{2.4}$$

denote a parametric family of conditional densities (with respect to some measure ξ_T) of Y_t given $Y_1, \ldots, Y_{t-1}$ and $X_1, \ldots, X_t$, evaluated at the rv's $Y_1, \ldots, Y_t$, $X_1, \ldots, X_t$. The conditional log–likelihood function of $\{Y_t : t \leq T\}$ given $\{X_t : t \leq T\}$ is $\Sigma_1^T \log f_t(\theta_1, \theta_3)$. The distribution of $\{X_t : t \leq T\}$ is assumed not to depend on θ_1 or θ_3. Let $\theta_{10} \in \Theta_1 \subset R^p$ and $\theta_{30} \in \Theta_3 \subset R^{p_3}$ denote the true values of θ_1 and θ_3, respectively, when no parameter instability or structural change occurs.

The ML estimator $\hat{\theta}(\pi) = (\hat{\theta}_1(\pi)', \hat{\theta}_2(\pi)', \hat{\theta}_3(\pi)')'$ is defined to minimize

$$-\Sigma_1^{T\pi} \log f_t(\theta_1, \theta_3) - \Sigma_{T\pi+1}^{T} \log f_t(\theta_2, \theta_3) \tag{2.5}$$

over $\theta = (\theta_1', \theta_2', \theta_3')' \in \Theta_1 \times \Theta_1 \times \Theta_3 = \Theta$. This estimator is a PSE estimator with

$$W_t = (Y_1', \ldots, Y_t', X_1', \ldots, X_t')' \in R^{k_t}, \ m_{1t}(\theta_1,\theta_3,\tau_1) = -\log f_t(\theta_1,\theta_3),$$

(2.6)

$$m_{3t}(\theta_1,\theta_3,\tau_1) = 0 \ \text{and} \ d(m,\gamma) = m_1 + m_2 + m_3 \ \text{for} \ m = (m_1,m_2,m_3)' \in R^3 .$$

No nuisance parameter estimators $\tilde{\tau}(\pi)$ or $\hat{\gamma}(\pi)$ arise in this example.

The choice of $m_t(\cdot,\cdot)$ and $d(\cdot,\cdot)$ in (2.6) is used to prove the consistency of the ML estimator $\hat{\theta}(\pi)$. To obtain its asymptotic distribution, however, it is more convenient to express $\hat{\theta}(\pi)$ as the solution to a set of first order conditions. That is, we take W_t as above and

$$m_{1t}(\theta_1, \theta_3, \tau_1) = -\frac{\partial}{\partial\theta_1}\log f_t(\theta_1, \theta_3), \ m_{3t}(\theta_1, \theta_3, \tau_1) = -\frac{\partial}{\partial\theta_3}\log f_t(\theta_1, \theta_3),$$

(2.7)

$$\text{and} \ d(m,\gamma) = m'm/2 .$$

If $\hat{\theta}(\pi)$ is consistent for some θ_0 and θ_0 is an interior point of Θ , then $\hat{\theta}(\pi)$ will satisfy the first order conditions from (2.6), and hence, will minimize (2.1) defined using (2.7), with probability that goes to one as $T \to \infty$. In this case, the same estimator $\hat{\theta}(\pi)$ can be expressed as a PSE estimator using (2.6) or (2.7), since the definition of a PSE estimator only requires (2.1) to be minimized with probability that goes to one as $T \to \infty$.

2.2. Uniform Consistency of Partial–Sample Extremum Estimators

Here we establish the consistency of $\hat{\theta}(\pi)$ for $\theta_0 = (\theta_{10}', \theta_{10}', \theta_{30}')'$ uniformly over $\pi \in \Pi$ for some $\theta_{10} \in \Theta_1$ and $\theta_{30} \in \Theta_3$. This result applies when no parameter instability or structural change occurs. The *uniformity* of the result is needed below to establish the weak convergence of $\hat{\theta}(\cdot)$ viewed as a process indexed by Π .

In what follows we avoid the devotion of space to questions of measurability by assuming the quantities dealt with are measurable or the probabilities dealt with are interpreted as inner or outer probabilities when necessary. Also, we pay little attention to the question of existence of a sequence of PSE estimators. Existence is assumed implicitly.

Sufficient conditions for existence include continuity of $d(\bar{m}_T(\theta,\pi,\hat{\underset{\sim}{\tau}}(\pi)), \hat{\gamma}(\pi))$ in θ with probability one and compactness of Θ. These conditions, of course, are far from necessary.

For consistency we assume the following:

ASSUMPTION 1: (a) $\sup_{\pi\in\Pi}\|\hat{\tau}_r(\pi) - \tau_0\| \xrightarrow{p} 0$ *for* $r = 1, 2$ *for some* $\tau_0 \in \mathcal{T}$ *and* $\sup_{\pi\in\Pi}\|\hat{\gamma}(\pi) - \gamma_0(\pi)\| \xrightarrow{p} 0$ *for some* $\gamma_0(\pi) \in \Gamma$ $\forall\pi \in \Pi$.

(b) $\{m_{rt}(\theta_1, \theta_3, \tau_1) : t \geq 1\}$ *satisfies a uniform* SLLN *over* $\Theta_1 \times \Theta_3 \times \mathcal{T}_0$, *where* $\mathcal{T}_0$ *is some neighborhood of* τ_0, *and* $m_r(\theta_1, \theta_3, \tau_1) = \lim_{T\to\infty} \frac{1}{T}\Sigma_1^T E m_{rt}(\theta_1, \theta_3, \tau_1)$ *exists uniformly over* $\Theta_1 \times \Theta_3 \times \mathcal{T}_0$ *for* $r = 1, 3$.[4]

(c) $m_r(\theta_1, \theta_3, \tau_1)$ *is continuous in* τ_1 *at* τ_0 *uniformly over* $(\theta_1, \theta_3) \in \Theta_1 \times \Theta_3$ *for* $r = 1, 3$ *and* $d(m,\gamma)$ *is uniformly continuous over* $\mathcal{M} \times \Gamma_0$, *where* Γ_0 $(\subset \Gamma)$ *contains an* ϵ*–neighborhood of* $\gamma_0(\pi)$ $\forall\pi \in \Pi$ *for some* $\epsilon > 0$, $\mathcal{M} = \{m \in R^v : m = m(\theta,\pi,\underset{\sim}{\tau})$ *for some* $\theta \in \Theta$, $\pi \in \Pi$, $\underset{\sim}{\tau} \in \mathcal{T}_0 \times \mathcal{T}_0\}$, *and* $m(\theta,\pi,\underset{\sim}{\tau}) = (\pi m_1(\theta_1,\theta_3,\tau_1)', (1-\pi)m_1(\theta_2,\theta_3,\tau_2)', \pi m_3(\theta_1,\theta_3,\tau_1)' + (1-\pi)m_3(\theta_2,\theta_3,\tau_2)')'$.

(d) *For every neighborhood* Θ_0 $(\subset \Theta)$ *of* θ_0, $\inf_{\pi\in\Pi}\Big[\inf_{\theta\in\Theta/\Theta_0} d(m(\theta, \pi, \underset{\sim}{\tau}_0), \gamma_0(\pi)) - d(m(\theta_0, \pi, \underset{\sim}{\tau}_0), \gamma_0(\pi))\Big] > 0$.

THEOREM 1: *Under Assumption* 1, *every sequence of extremum estimators* $\{\hat{\theta}(\cdot)\}$ *satisfies* $\sup_{\pi\in\Pi}\|\hat{\theta}(\pi) - \theta_0\| \xrightarrow{p} 0$.

We now discuss Assumption 1. Assumption 1(a) either holds trivially because no preliminary estimators $\hat{\underset{\sim}{\tau}}(\pi)$ and $\hat{\gamma}(\pi)$ appear or it can be verified by the application of Theorem 1 to $\hat{\underset{\sim}{\tau}}(\pi)$ and $\hat{\gamma}(\pi)$ rather than $\hat{\theta}(\pi)$. Assumption 1(b) can be verified under broad assumptions regarding temporal dependence and non–identical distributions using the uniform strong laws of large numbers (SLLN) provided, for example, by Andrews (1987, 1989d), Pötscher and Prucha (1989), and Newey (1989). Assumption 1(c) is usually straightforward to verify and holds in most applications. Assumption 1(d) is the uniqueness assumption that ensures that $\{\hat{\theta}(\pi)\}$ converges to a point θ_0 rather than to a multi–element subset of Θ .

EXAMPLE 1 (GMM, cont.): Sufficient conditions (similar to those of Hansen (1982)) for the GMM estimator to satisfy Assumption 1 are given by

ASSUMPTION GMM–1: (a) $\sup_{\pi\in\Pi}\|a_{1,T\pi}-a_0\| + \sup_{\pi\in\Pi}\|a_{2,T\pi}-a_0\| \xrightarrow{p} 0$ *for some* $p \times v_1$ *matrix* a_0 .

(b) $\{W_t : -\infty < t < \infty\}$ *is stationary and ergodic.*

(c) $f(w, \theta_1)$ *is continuous in* θ_1 *uniformly over* $\theta_1 \in \Theta_1$ *for each* w *in the support of* W_t *and* $E \sup_{\theta_1\in\Theta_1} \|f(W_t, \theta_1)\| < \infty$.

(d) Θ_1 *is bounded.*

(e) $Ef(W_t, \theta_{10}) = \underset{\sim}{0}$ *and* $\inf_{\theta_1\in\Theta_1/\Theta_{10}} Ef(W_t, \theta_1)'a_0'a_0Ef(W_t, \theta_1) > 0$ *for all neighborhoods* Θ_{10} *of* θ_{10} .

To see that Assumption GMM–1 implies Assumption 1, note that GMM–1(a) implies 1(a), GMM–1(b), (c), and (d) imply 1(b) using Andrews' (1989d, Thm. 6 and As. TSE–1C) uniform SLLN, GMM–1(c) implies 1(c) (since $m(\theta,\pi,\underset{\sim}{\tau}) = (\pi Ef(W_t, \theta_1)', (1-\pi)Ef(W_t, \theta_2)')'$ and GMM–1(c) implies that $\sup_{\theta_1\in\Theta_1} Ef(W_t, \theta_1) < \infty$), and GMM–1(e) implies 1(d). Assumption GMM–1(a) can be made more primitive if $a_{1,T\pi}$ and $a_{2,T\pi}$ are defined more explicitly.

EXAMPLE 2 (ML, cont.): For brevity, we do not state an analogue of Assumption GMM–1 under which Assumption 1 holds for ML estimators. Instead, we note that with $m_{1t}(\cdot,\cdot)$ and $d(\cdot,\cdot)$ as in (2.6), Assumption 1(a) holds trivially, 1(b) requires that $\{-\log f_t(\theta_1, \theta_3) : t \geq 1\}$ satisfies a uniform SLLN over $\Theta_1 \times \Theta_3$ and that $m_1(\theta_1, \theta_3) = \lim_{T\to\infty} -\frac{1}{T}\Sigma_1^T E \log f_t(\theta_1, \theta_3)$ exists uniformly over $\Theta_1 \times \Theta_3$, 1(c) holds automatically, and 1(d) requires that $\inf_{(\theta_1,\theta_3)\in\Theta_1\times\Theta_3/\Theta_{10}\times\Theta_{30}} (m_1(\theta_1, \theta_3) - m_1(\theta_{10}, \theta_{30})) > 0$ for all neighborhoods Θ_{10} and Θ_{30} of θ_{10} and θ_{30}. These conditions are just the same as the conditions used to prove strong consistency of the ML estimator when the full sample is used.

2.3. Weak Convergence of Partial-Sample Extremum Estimators

The asymptotic distribution of the PSE estimator $\hat{\theta}(\cdot)$ depends on the following matrices:

$$\Sigma = \begin{bmatrix} S & S_{13} \\ S_{13}' & S_{33} \end{bmatrix} = \lim_{T\to\infty} \mathrm{Var}\left[\frac{1}{\sqrt{T}}\Sigma_1^T \begin{bmatrix} m_{1t}(\theta_{10}, \theta_{30}, \tau_0) \\ m_{3t}(\theta_{10}, \theta_{30}, \tau_0) \end{bmatrix}\right],$$

$$(2.8) \qquad \begin{bmatrix} M & M_{13} \\ M_{31} & M_{33} \end{bmatrix} = \lim_{T\to\infty} \frac{1}{T}\Sigma_1^T E\frac{\partial}{\partial(\theta_1', \theta_3')} \begin{bmatrix} m_{1t}(\theta_{10}, \theta_{30}, \tau_0) \\ m_{3t}(\theta_{10}, \theta_{30}, \tau_0) \end{bmatrix},$$

$$M(\pi) = \lim_{T\to\infty} E\frac{\partial}{\partial\theta'}\bar{m}_T(\theta_0, \pi, \underset{\sim}{\tau}_0) = \begin{bmatrix} \pi M & \underset{\sim}{0} & \pi M_{13} \\ \underset{\sim}{0} & (1-\pi)M & (1-\pi)M_{13} \\ \pi M_{31} & (1-\pi)M_{31} & M_{33} \end{bmatrix}, \text{ and}$$

$$D(\pi) = \frac{\partial^2}{\partial m\,\partial m'} d(m, \gamma_0(\pi))\Big|_{m=m(\theta_0,\pi,\underset{\sim}{\tau}_0)},$$

where $m(\theta_0, \pi, \underset{\sim}{\tau}_0) = \lim_{T\to\infty} E\bar{m}_T(\theta_0, \pi, \underset{\sim}{\tau}_0)$. As noted above (see Example 2), the functions $m_{rt}(\cdot,\cdot,\cdot)$ and $d(\cdot,\cdot)$ may be chosen differently in this section and in Section 2.2.

EXAMPLE 1 (GMM, cont.): In this example, we have

$$(2.9) \qquad \begin{gathered} S = \lim_{T\to\infty} \mathrm{Var}\left[\frac{1}{\sqrt{T}}\Sigma_1^T f(W_t, \theta_{10})\right], \quad M = \lim_{T\to\infty} \frac{1}{T}\Sigma_1^T E\frac{\partial}{\partial\theta_1'} f(W_t, \theta_{10}), \\ M(\pi) = \begin{bmatrix} \pi M & \underset{\sim}{0} \\ \underset{\sim}{0} & (1-\pi)M \end{bmatrix}, \text{ and } D(\pi) = \begin{bmatrix} a_0'a_0/\pi & \underset{\sim}{0} \\ \underset{\sim}{0} & a_0'a_0/(1-\pi) \end{bmatrix}. \end{gathered}$$

EXAMPLE 2 (ML, cont.): For the weak convergence results of this section, we use the definition of $m_{rt}(\cdot,\cdot)$ and $d(\cdot,\cdot)$ given in (2.7) for the ML estimator. This yields

$$S = \lim_{T\to\infty} \operatorname{Var}\left[\frac{1}{\sqrt{T}}\Sigma_1^T \frac{\partial}{\partial\theta_1}\log f_t(\theta_{10},\theta_{30})\right], \quad M = \lim_{T\to\infty} -\frac{1}{T}\Sigma_1^T E\frac{\partial^2}{\partial\theta_1\partial\theta_1'}\log f_t(\theta_{10},\theta_{30}),$$

$$M_{13} = M_{31}' = \lim_{T\to\infty} -\frac{1}{T}\Sigma_1^T E\frac{\partial^2}{\partial\theta_1\partial\theta_3'}\log f_t(\theta_{10},\theta_{30}),$$

$$M_{33} = \lim_{T\to\infty} -\frac{1}{T}\Sigma_1^T E\frac{\partial^2}{\partial\theta_3\partial\theta_3'}\log f_t(\theta_{10},\theta_{30}), \tag{2.10}$$

$$M(\pi) = \begin{bmatrix} \pi M & \underset{\sim}{0} & \pi M_{13} \\ \underset{\sim}{0} & (1-\pi)M & (1-\pi)M_{13} \\ \pi M_{13}' & (1-\pi)M_{13}' & M_{33} \end{bmatrix}, \text{ and } D(\pi) = I_{2p+p_3}.$$

Let $\{(B_1(\pi)', B_3(\pi)')' : \pi \in [0,1]\}$ denote a v_1+v_3–vector of independent Brownian motions on [0,1]. Let $o_{p\pi}(1)$ and $O_{p\pi}(1)$ denote quantities that are $o_p(1)$ and $O_p(1)$ uniformly over $\pi \in \Pi$ respectively. Let Θ_{10}, Θ_{30}, and $\mathcal{T}_0$ denote some neighborhoods of θ_{10}, θ_{30}, and τ_0 respectively.

The following assumption is sufficient to obtain the weak convergence of the PSE estimator $\hat\theta(\cdot)$ as a process indexed by $\pi \in \Pi$ to a function of the Brownian motion vector $(B_1(\cdot)', B_3(\cdot)')'$:

ASSUMPTION 2: (a) $\sup_{\pi\in\Pi}\|\hat\theta(\pi) - \theta_0\| \xrightarrow{p} 0$ *for some* $\theta_0 = (\theta_{10}', \theta_{20}', \theta_{30}')'$ *that is in the interior of* Θ *and satisfies* $\theta_{10} = \theta_{20}$.

(b) $\sup_{\pi\in\Pi}\|\hat\gamma(\pi) - \gamma_0(\pi)\| \xrightarrow{p} 0$ *for some* $\gamma_0(\pi) \in \Gamma$ $\forall \pi \in \Pi$.

(c) $\sqrt{T}(\hat\tau_r(\pi) - \tau_0) = O_{p\pi}(1)$ *for some* $\tau_0 \in \mathcal{T}$ *for* $r = 1, 2$.

(d) $\sup_{\pi\in\Pi}\left\|\frac{\partial}{\partial m}d(E\bar m_T(\theta_0, \pi, \underset{\sim}{\tau}_0), \hat\gamma(\pi))\right\| = o_p(T^{-1/2})$, *where* $\underset{\sim}{\tau}_0 = (\tau_0', \tau_0')'$.

(e) $\{(m_{1t}(\theta_{10},\theta_{30},\tau_0)', m_{3t}(\theta_{10},\theta_{30},\tau_0)')' : t \geq 1\}$ *satisfies an invariance principle with covariance matrix* Σ. *That is, for* $\nu_{rT}(\pi) = \frac{1}{\sqrt{T}}\Sigma_1^{T\pi}(m_{rt}(\theta_{10},\theta_{30},\tau_0) - Em_{rt}(\theta_{10},\theta_{30},\tau_0))$,

$\mathrm{r} = 1, 3$, *we have* $\nu_{\mathrm{T}}(\cdot) = \begin{bmatrix} \nu_{1\mathrm{T}}(\cdot) \\ \nu_{3\mathrm{T}}(\cdot) \end{bmatrix} \Rightarrow \begin{bmatrix} \nu_1(\cdot) \\ \nu_3(\cdot) \end{bmatrix} = \begin{bmatrix} \mathrm{S}^{1/2} & \underset{\sim}{0} \\ \mathrm{S}_{31}^0 & \mathrm{S}_{33}^0 \end{bmatrix} \begin{bmatrix} \mathrm{B}_1(\cdot) \\ \mathrm{B}_3(\cdot) \end{bmatrix}$ *as a process indexed by* $\pi \in [0,1]$, *where* $\mathrm{S}_{31}^0 \in \mathrm{R}^{\mathrm{v}_3 \times \mathrm{v}_1}$ *and* $\mathrm{S}_{33}^0 \in \mathrm{R}^{\mathrm{v}_3 \times \mathrm{v}_3}$ *are such that* $\begin{bmatrix} \mathrm{S}^{1/2} & \underset{\sim}{0} \\ \mathrm{S}_{31}^0 & \mathrm{S}_{33}^0 \end{bmatrix} \begin{bmatrix} \mathrm{S}^{1/2} & \underset{\sim}{0} \\ \mathrm{S}_{31}^0 & \mathrm{S}_{33}^0 \end{bmatrix}' = \Sigma$.

(f) $\frac{\partial}{\partial \mathrm{m}} \mathrm{d}(\mathrm{m},\gamma)$ *and* $\frac{\partial^2}{\partial \mathrm{m} \partial \mathrm{m}'} \mathrm{d}(\mathrm{m},\gamma)$ *exist for* $(\mathrm{m},\gamma) \in \mathcal{M}_0 \times \Gamma_0$ *and are uniformly continuous on* $\{(\mathrm{m}(\theta_0,\pi,\underset{\sim}{\tau}_0), \gamma_0(\pi)) : \pi \in \Pi\}$, *where* $\mathcal{M}_0$ *and* Γ_0 *contain* ϵ*–neighborhoods of* $\mathrm{m}(\theta_0, \pi, \underset{\sim}{\tau}_0)$ *and* $\gamma_0(\pi)$, *respectively, for all* $\pi \in \Pi$ *and some* $\epsilon > 0$ *and* $\mathrm{m}(\theta_0, \pi, \underset{\sim}{\tau}_0) = \lim_{\mathrm{T} \to \infty} \mathrm{E}\bar{\mathrm{m}}_{\mathrm{T}}(\theta_0, \pi, \underset{\sim}{\tau}_0)$.

(g) $\mathrm{m}_{\mathrm{rt}}(\theta_1, \theta_3, \tau_1)$ *is twice, twice, and once continuously differentiable in* θ_1, θ_3, *and* τ_1, *respectively, on* $\Theta_{10} \times \Theta_{30} \times \mathcal{T}_0$ $\forall \mathrm{t} \geq 1$, $\forall \omega \in \Omega$, *for* $\mathrm{r} = 1, 3$. *The sequences* $\{\mathrm{m}_{\mathrm{rt}}(\theta_1, \theta_3, \tau_1)\}$, $\left\{\frac{\partial}{\partial(\theta_1', \theta_3')} \mathrm{m}_{\mathrm{rt}}(\theta_1, \theta_3, \tau_1)\right\}$, *and* $\left\{\frac{\partial}{\partial \tau_1'} \mathrm{m}_{\mathrm{rt}}(\theta_1, \theta_3, \tau_1)\right\}$ *satisfy uniform* SLLN's *over* $\Theta_{10} \times \Theta_{30} \times \mathcal{T}_0$ *for* $\mathrm{r} = 1, 3$. *The limits* $\mathrm{m}_{\mathrm{r}}(\theta_1, \theta_3, \tau_1) = \lim_{\mathrm{T} \to \infty} \frac{1}{\mathrm{T}} \Sigma_1^{\mathrm{T}} \mathrm{E}\mathrm{m}_{\mathrm{rt}}(\theta_1, \theta_3, \tau_1)$, $\lim_{\mathrm{T} \to \infty} \frac{1}{\mathrm{T}} \Sigma_1^{\mathrm{T}} \mathrm{E} \frac{\partial}{\partial(\theta_1', \theta_3')} \mathrm{m}_{\mathrm{rt}}(\theta_1, \theta_3, \tau_1)$, *and* $\mathrm{dm}_{\mathrm{r}}(\theta_1, \theta_3, \tau_1) = \lim_{\mathrm{T} \to \infty} \frac{1}{\mathrm{T}} \Sigma_1^{\mathrm{T}} \mathrm{E} \frac{\partial}{\partial \tau_1'} \mathrm{m}_{\mathrm{rt}}(\theta_1, \theta_3, \tau_1)$ *exist uniformly over* $\Theta_{10} \times \Theta_{30} \times \mathcal{T}_0$ *and are continuous at* $(\theta_{10}, \theta_{30}, \tau_0)$ *for* $\mathrm{r} = 1, 3$. $\mathrm{dm}_{\mathrm{r}}(\theta_{10}, \theta_{30}, \tau_0) = \underset{\sim}{0}$ *for* $\mathrm{r} = 1, 3$. *In addition,* $\overline{\lim_{\mathrm{T} \to \infty}} \frac{1}{\mathrm{T}} \Sigma_1^{\mathrm{T}} \mathrm{E} \sup_{\theta_1 \in \Theta_{10}, \theta_3 \in \Theta_{30}, \tau_1 \in \mathcal{T}_0} \left\| \frac{\partial^2}{\partial(\theta_1', \theta_3')' \partial(\theta_1', \theta_3')} \mathrm{m}_{\mathrm{rt}}(\theta_1, \theta_3, \tau_1) \right\| < \infty$ *for* $\mathrm{r} = 1, 3$.

(h) $\mathrm{M}(\pi)'\mathrm{D}(\pi)\mathrm{M}(\pi)$ *is nonsingular* $\forall \pi \in \Pi$ *and has eigenvalues bounded away from zero.*

THEOREM 2: *Under Assumption* 2, *every sequence of partial–sample extremum estimators* $\{\hat{\theta}(\cdot)\}$ *satisfies*

$$\sqrt{\mathrm{T}}(\hat{\theta}(\cdot) - \theta_0) \Rightarrow -(\mathrm{M}(\cdot)'\mathrm{D}(\cdot)\mathrm{M}(\cdot))^{-1}\mathrm{M}(\cdot)'\mathrm{D}(\cdot)\mathrm{G}(\cdot)$$

as a process indexed by $\pi \in \Pi$, *where* Π *has closure contained in* (0,1), $\mathrm{G}(\cdot) = (\nu_1(\cdot)', \nu_1(1)' - \nu_1(\cdot)', \nu_3(1)')'$, *and* $(\nu_1(\cdot)', \nu_3(\cdot)')'$ *is the Gaussian process defined in Assumption* 2(e).

COMMENTS: 1. In the case of pure structural change, the result of Theorem 2 can be expressed as

$$(2.11) \qquad \sqrt{T}\begin{bmatrix} \hat{\theta}_1(\cdot) - \theta_{10} \\ \hat{\theta}_2(\cdot) - \theta_{10} \end{bmatrix} \Rightarrow \begin{bmatrix} \frac{1}{\iota(\cdot)}(M'D_1(\cdot)M)^{-1}M'D_1(\cdot)S^{1/2}B_1(\cdot) \\ \frac{1}{1-\iota(\cdot)}(M'D_2(\cdot)M)^{-1}M'D_2(\cdot)S^{1/2}(B_1(1) - B_1(\cdot)) \end{bmatrix},$$

where $\iota(\cdot)$ is the identity function on Π, i.e. $\iota(\pi) = \pi \ \forall \pi \in \Pi$, and $D_1(\cdot)$ and $D_2(\cdot)$ are the upper and lower $v_1 \times v_1$ blocks of $D(\cdot)$ respectively.

2. In the case where $2p + p_3 = v$ (i.e., the dimension of θ equals the dimension of $\bar{m}_T(\theta,\pi,\underset{\sim}{\tau})$), the limit process of Theorem 2 simplifies to $-M(\cdot)^{-1}G(\cdot)$, since $(M(\cdot)'D(\cdot)M(\cdot))^{-1} = M(\cdot)^{-1}D(\cdot)^{-1}(M(\cdot)')^{-1}$.

Next we discuss Assumption 2. Assumption 2(a) can be established by Theorem 1. Assumptions 2(b) and (c) hold trivially if no preliminary estimators $\hat{\gamma}(\pi)$ and $\hat{\underset{\sim}{\tau}}(\pi)$ appear. Otherwise, they can be established by applying Theorems 1 and 2 to $\hat{\gamma}(\cdot)$ and $\hat{\underset{\sim}{\tau}}(\cdot)$ respectively. Assumption 2(d) usually is simple to verify since $Em_{rt}(\theta_{10}, \theta_{30}, \tau_0)$ usually equals zero $\forall t \geq 1$, for $r = 1, 3$.

Assumption 2(e) can be verified by applying a multivariate invariance principle for dnid rv's (e.g., see Phillips and Durlauf (1986, Thm. 2.1) and Eberlein (1986)) or by showing that the rv's $\{(m_{1t}(\theta_{10}, \theta_{30}, \tau_0)', m_{3t}(\theta_{10}, \theta_{30}, \tau_0)')\alpha : t \geq 1\}$ satisfy a univariate invariance principle with variance $\alpha'\Sigma\alpha$ for each unit v_1+v_3–vector α and that $\{\nu_T(\cdot) : T \geq 1\}$ has asymptotically independent increments.[5] For example, univariate invariance principles are given for dnid rv's by McLeish (1975, Thms. 3.8 and 4.2; 1977, Thm. 2.4 and Cor. 2.11), Herrndorf (1984, Thm. and Cors. 1–4), and Wooldridge and White (1988, Thm. 2.11 and Cors. 3.1–3.2). Note that the property of asymptotically independent increments of $\{\nu_T(\cdot) : T \geq 1\}$ generally is implied by the weak dependence conditions on $\{m_{rt}(\theta_{10}, \theta_{30}, \tau_0) : t \geq 1\}$ that are used to obtain the univariate invariance principle for $\{(m_{1t}(\theta_{10}, \theta_{30}, \tau_0)', m_{3t}(\theta_{10}, \theta_{30}, \tau_0)')\alpha : t \geq 1\}$ for arbitrary unit v_1+v_3–vector α.

Since $d(m,\gamma)$ is often of the form $m'\gamma m/2$, Assumption 2(f) is not restrictive and is easy to verify. Assumption 2(g) is a standard requirement of smoothness of $m_{rt}(\theta_1, \theta_3, \tau_1)$ in θ_1, θ_3, and τ_1, the existence of certain limiting averages of expectations, and non-explosive non-trending behavior of the summands $\{m_{rt}(\theta_1, \theta_3, \tau_1)\}$ and their first two derivatives. The uniform SLLN's referenced in Section 2.2 above can be used in verifying this assumption. The smoothness assumptions could be avoided by using the approach taken in Andrews (1989a,b,c). In addition, using the latter approach, the preliminary nuisance parameter estimator $\hat{\underset{\sim}{\tau}}(\pi)$ could be infinite dimensional for each π. Assumption 2(g) also imposes an asymptotic orthogonality condition between the estimators of $(\theta_{10}', \theta_{30}')'$ and τ_0. This condition is easy to verify when it holds.

Assumption 2(h) imposes a nonsingularity condition that ensures that the estimator $\hat{\theta}(\pi)$ has a nonsingular asymptotic variance $\forall \pi \in \Pi$. In many examples, this assumption is satisfied if the covariance matrix Σ of Assumption 2(e) is nonsingular, see Examples 1 and 2 below.

EXAMPLE 1 (GMM, cont.): In the GMM example, the limit process of $\sqrt{T}(\hat{\theta}(\cdot) - \theta_0)$ given in Theorem 2 is

$$\left[\begin{array}{c} \frac{1}{\iota(\cdot)}(M'a_0'a_0M)^{-1}M'a_0'a_0S^{1/2}\, B_1(\cdot) \\ \frac{1}{1-\iota(\cdot)}(M'a_0'a_0M)^{-1}M'a_0'a_0S^{1/2}(B_1(1) - B_1(\cdot)) \end{array}\right], \tag{2.12}$$

where M and S are as in (2.9), $a_0'a_0$ is as in Assumption GMM–1, and $\iota(\pi) = \pi$. Also, Assumption GMM–1 plus the following Assumption GMM–2 are sufficient for Assumption 2 (where McLeish's (1975) Thm. 2.5 is used to establish Assumption 2(e)). Let f_t abbreviate $f(W_t, \theta_{10})$, let $\mathcal{F}_m = \sigma(\{W_t : -\infty < t \le m\})$, and let $\|\cdot\|_2$ denote the L^2 norm.

ASSUMPTION GMM–2: (i) *θ_{10} is in the interior of Θ_1.*

(ii) $Ef_t'f_t < \infty$ *and* $\mathrm{Var}\left[\frac{1}{\sqrt{T}}\Sigma_1^T f_t\right] \to S$ *for some positive definite* $v_1 \times v_1$ *matrix* S.

(iii) $E\left|\mathrm{Var}\left[\frac{1}{\sqrt{T}}\Sigma_1^T h'f_t | \mathcal{F}_{-m}\right] - h'Sh\right| \to 0$ *as* $\min\{T,m\} \to \infty$ $\forall h \in R^{v_1}$.

(iv) $\psi_k = o(k^{-\xi})$ *for some* $\xi > 1/2$, *where* $\psi_k = \sup_{h \in R^{v_1}, \|h\|=1} \|E(h'f_t | \mathcal{F}_{t-k})\|_2$.

(v) $f(W_t, \theta_1)$ *is twice continuously differentiable in* θ_1 *on some neighborhood* Θ_{10} *of* θ_{10} *for all realizations of* W_t, $E \sup_{\theta_1 \in \Theta_{10}} \left\|\frac{\partial}{\partial \theta_1'} f(W_t, \theta_1)\right\| < \infty$, *and*

$E \sup_{\theta_1 \in \Theta_{10}} \left\|\frac{\partial^2}{\partial \theta_1 \partial \theta_1'} f(W_t, \theta_1)\right\| < \infty$.

(vi) $E\frac{\partial}{\partial \theta_1'} f(W_t, \theta_{10})$ *is full rank* p $(\le v_1)$ *and* $a_0'a_0$ *is nonsingular.*

More primitive conditions than GMM–2(iii) and (iv) can be given in terms of strong mixing or near epoch dependent rv's using Cor. 3.9 and Thm. 4.2 of McLeish (1975). For example, GMM–2(iii) and (iv) can be replaced by: $\{W_t\}$ is strong mixing with strong mixing coefficients $\{\alpha(n)\}$, $\sum_{n=1}^{\infty} \alpha(n)^{1-2/\beta} < \infty$ and $E\|f_t\|^\beta < \infty$ for some $\beta > 2$, and $Ef_t = \underset{\sim}{0}$.

EXAMPLE 2 (ML, cont.): For ML estimators, the limit process of Theorem 2 equals

$$(2.13) \qquad \begin{bmatrix} \iota(\cdot)M & \underset{\sim}{0} & \iota(\cdot)M_{13} \\ \underset{\sim}{0} & (1-\iota(\cdot))M & (1-\iota(\cdot))M_{13} \\ \iota(\cdot)M_{13}' & (1-\iota(\cdot))M_{13}' & M_{33} \end{bmatrix}^{-1} \begin{bmatrix} \nu_1(\cdot) \\ \nu_1(1) - \nu_1(\cdot) \\ \nu_3(\cdot) \end{bmatrix},$$

since $D(\cdot) = I_{2p+p_3}$ and $M(\cdot)$ is symmetric.

In this example, Assumption 2(a) holds by Theorem 1 under the conditions outlined above. Assumptions 2(b) and (c) hold trivially, since no preliminary estimators $\hat{\gamma}(\pi)$ and $\hat{\underset{\sim}{\tau}}(\pi)$ arise. Assumption 2(d) holds since $E\bar{m}_T(\theta_0, \pi, \underset{\sim}{\tau}_0) = \underset{\sim}{0}$ provided the conditional density $f_t(\theta_{10}, \theta_{30})$ is sufficiently regular to permit interchange of the $\frac{\partial}{\partial(\theta_1', \theta_3')'}$ and

$\int$ operations in the expression $\frac{\partial}{\partial(\theta_1', \theta_3')'}\int f_t(\theta_{10}, \theta_{30})d\xi_T$. Assumption 2(e) requires $\left\{\frac{\partial}{\partial(\theta_1', \theta_3')'}\log f_t(\theta_{10}, \theta_{30}) : t \geq 1\right\}$ to satisfy an invariance principle. Any of the results referenced above can be used for this. Assumption 2(f) holds trivially. Assumption 2(g) holds if (i) $\log f_t(\theta_1, \theta_3)$ is three times continuously differentiable on $\Theta_{10} \times \Theta_{30}$ for all realizations of $\{W_{Tt}\}$, (ii) $\left\{\frac{\partial}{\partial(\theta_1', \theta_3')'}\log f_t(\theta_1, \theta_3)\right\}$ and $\left\{\frac{\partial^2}{\partial(\theta_1', \theta_3')'\partial(\theta_1', \theta_3')}\log f_t(\theta_1, \theta_3)\right\}$ satisfy uniform SLLN's over $\Theta_{10} \times \Theta_{30}$, (iii) $\lim_{T\to\infty} \frac{1}{T}\Sigma_1^T E\frac{\partial}{\partial(\theta_1', \theta_3')'}\log f_t(\theta_1, \theta_3)$ and $\lim_{T\to\infty} \frac{1}{T}\Sigma_1^T E\frac{\partial^2}{\partial(\theta_1', \theta_3')'\partial(\theta_1', \theta_3')}\log f_t(\theta_1, \theta_3)$ exist uniformly over $\Theta_{10} \times \Theta_{30}$ (as occurs, e.g., in the stationary q–th order Markov case), and (iv) $\overline{\lim_{T\to\infty}} \frac{1}{T}\Sigma_1^T E \sup_{\theta_1 \in \Theta_{10}, \theta_3 \in \Theta_{30}} \left\|\frac{\partial}{\partial\theta_r'} \operatorname{vec} \frac{\partial^2}{\partial\theta_s \partial\theta_q'}\log f_t(\theta_1, \theta_3)\right\| < \infty$ $\forall r, s, q = 1, 3$. Assumption 2(h) holds provided the asymptotic information matrix $-\lim_{T\to\infty} \frac{1}{T}\Sigma_1^T E\frac{\partial^2}{\partial(\theta_1', \theta_3')'\partial(\theta_1', \theta_3')}\log f_t(\theta_{10}, \theta_{30})$ is nonsingular. (The latter condition is sufficient for Assumption 2(h), because a formula for the determinant of a partitioned matrix shows that the determinant of $M(\pi)$ equals $\pi(1-\pi)$ times the determinant of the upper $p \times p$ block of the asymptotic information matrix times the determinant of the asymptotic information matrix.)

2.4. Covariance Matrix Estimation for Partial–Sample Extremum Estimators

The Wald statistic defined in Section 3 below is based on the vector $\sqrt{T}(\hat{\theta}_1(\cdot) - \hat{\theta}_2(\cdot))$. Here we consider estimation of the unknown constants that appear in the limit distribution of $\sqrt{T}(\hat{\theta}_1(\cdot) - \hat{\theta}_2(\cdot))$.

For $m = (m_1', m_2', m_3')' \in R^{v_1+v_1+v_3}$, let

$$D_r(\pi) = \frac{\partial^2}{\partial m_r \partial m_r'}d(m, \gamma_0(\pi))\Big|_{m=m(\theta_0,\pi,\underset{\sim}{\gamma}_0)}, \quad \delta_r(\pi) = \begin{cases} \pi & \text{for } r = 1 \\ 1-\pi & \text{for } r = 2 \end{cases}, \text{ and}$$

$$V_r(\pi) = \delta_r^{-1}(\pi)(M'D_r(\pi)M)^{-1}M'D_r(\pi)SD_r(\pi)M(M'D_r(\pi)M)^{-1} \text{ for } r = 1, 2. \tag{2.14}$$

For fixed $\pi \in \Pi$, $V_1(\pi) + V_2(\pi)$ is the asymptotic variance of $\sqrt{T}(\hat{\theta}_1(\pi) - \hat{\theta}_2(\pi))$ under Assumption 2 and an additional assumption given in Section 3. Note that if $p = v_1$ (i.e. the dimension of θ_1 equals the dimension of $m_{1t}(\theta_1, \theta_3, \tau_1)$), $V_r(\pi)$ simplifies to $\delta_r^{-1}(\pi)M^{-1}S(M^{-1})'$, since $(M'D_r(\pi)M)^{-1} = M^{-1}D_r(\pi)^{-1}(M')^{-1}$. In this case, the estimators introduced below can be simplified accordingly.

Estimators of $\hat{V}_1(\pi)$ and $\hat{V}_2(\pi)$ are given by

$$\hat{V}_r = \delta_r^{-1}(\hat{M}_r'\hat{D}_r\hat{M}_r)^{-1}\hat{M}_r'\hat{D}_r\hat{S}_r\hat{D}_r\hat{M}_r(\hat{M}_r'\hat{D}_r\hat{M}_r)^{-1} \text{ for } r = 1, 2, \tag{2.15}$$ [6]

where the dependence of $\hat{V}_r$, $\hat{M}_r$, $\hat{D}_r$, δ_r, and $\hat{S}_r$ on π is suppressed for notational simplicity. The latter matrices can be defined in two ways. The first way uses only the data for $t = 1, \ldots, T\pi$ for the case $r = 1$ and only the data for $t = T\pi+1, \ldots, T$ for the case $r = 2$:

$$\begin{aligned} \hat{M}_1(\pi) &= \frac{1}{T\pi}\Sigma_1^{T\pi}\frac{\partial}{\partial\theta_1'}m_{1t}(\hat{\theta}_1(\pi), \hat{\theta}_3(\pi), \hat{\tau}_1(\pi)), \\ \hat{M}_2(\pi) &= \frac{1}{T-T\pi}\Sigma_{T\pi+1}^{T}\frac{\partial}{\partial\theta_2'}m_{1t}(\hat{\theta}_2(\pi), \hat{\theta}_3(\pi), \hat{\tau}_2(\pi)), \\ \hat{D}_r(\pi) &= \frac{\partial^2}{\partial m_r \partial m_r'}d(\bar{m}_T(\hat{\theta}(\pi), \pi, \hat{\underset{\sim}{\tau}}(\pi)), \hat{\gamma}(\pi)), \end{aligned} \tag{2.16}$$

and $\hat{S}_r(\pi)$ is as defined below, for $r = 1, 2$. The second way uses all of the data for $r = 1$ and $r = 2$:

$$\begin{aligned} \hat{M}_r(\pi) &= \hat{M} = \frac{1}{T}\Sigma_1^T\frac{\partial}{\partial\theta_1'}m_{1t}(\hat{\theta}_1(1), \hat{\theta}_3(1), \hat{\tau}_1(1)) \text{ and} \\ \hat{D}_r(\pi) &= \frac{\partial^2}{\partial m_r \partial m_r'}d(\bar{m}_T(\hat{\theta}(1), \pi, \hat{\underset{\sim}{\tau}}(1)), \hat{\gamma}(\pi)). \end{aligned} \tag{2.17}$$

The estimators defined in (2.16) and (2.17) have the same probability limits under the null hypothesis (using Assumption 2) and under sequences of local alternatives (see Section 4). They do not necessarily have the same probability limits, however, under sequences of fixed alternatives. One might argue that the estimators of (2.16) are preferable to those of (2.17) in terms of the power of Wald tests based upon them, but vice versa in terms of the closeness of the true and nominal sizes of such tests.

The $v_1 \times v_1$ matrix S equals $\lim_{T\to\infty} \mathrm{Var}\left[\frac{1}{\sqrt{T}}\Sigma_1^T m_{1t}(\theta_{10}, \theta_{30}, \tau_0)\right]$, see (2.8). If $\{m_{1t}(\theta_{10}, \theta_{30}, \tau_0)\}$ is a sequence of mean zero, uncorrelated rv's, then $S = \lim_{T\to\infty} \frac{1}{T}\Sigma_1^T E m_{1t}(\theta_{10}, \theta_{30}, \tau_0) m_{1t}(\theta_{10}, \theta_{30}, \tau_0)'$. In this case, the estimators $\hat{S}_r(\pi)$ corresponding to equations (2.16) and (2.17) can be taken to be

$$\text{(2.18)} \quad \begin{aligned} \hat{S}_1(\pi) &= \frac{1}{T\pi}\Sigma_1^{T\pi} m_{1t}(\hat{\theta}_1(\pi), \hat{\theta}_3(\pi), \hat{\tau}_1(\pi)) m_{1t}(\hat{\theta}_1(\pi), \hat{\theta}_3(\pi), \hat{\tau}_1(\pi))', \\ \hat{S}_2(\pi) &= \frac{1}{T-T\pi}\Sigma_{T\pi+1}^{T} m_{1t}(\hat{\theta}_2(\pi), \hat{\theta}_3(\pi), \hat{\tau}_2(\pi)) m_{1t}(\hat{\theta}_2(\pi), \hat{\theta}_3(\pi), \hat{\tau}_2(\pi))', \text{ and} \end{aligned}$$

$$\text{(2.19)} \quad \hat{S}_r(\pi) = \hat{S} = \frac{1}{T}\Sigma_1^T m_{1t}(\hat{\theta}_1(1), \hat{\theta}_3(1), \hat{\tau}_1(1)) m_{1t}(\hat{\theta}_1(1), \hat{\theta}_3(1), \hat{\tau}_1(1))' \text{ for } r = 1, 2,$$

respectively.

Alternatively, if $\{m_{1t}(\theta_{10}, \theta_{30}, \tau_0)\}$ is a sequence of mean zero, temporally dependent rv's, then

$$\text{(2.20)} \quad \begin{aligned} S = &\sum_{v=0}^{\infty} \lim_{T\to\infty} \frac{1}{T}\Sigma_{v+1}^T E m_{1t}(\theta_{10}, \theta_{30}, \tau_0) m_{1t-v}(\theta_{10}, \theta_{30}, \tau_0)' \\ &+ \sum_{v=1}^{\infty} \lim_{T\to\infty} \frac{1}{T}\Sigma_{v+1}^T E m_{1t-v}(\theta_{10}, \theta_{30}, \tau_0) m_{1t}(\theta_{10}, \theta_{30}, \tau_0)'. \end{aligned}$$

In this case, the estimator $\hat{S}_r(\pi)$ corresponding to (2.17) can be taken to be

$$\text{(2.21)} \quad \begin{aligned} \hat{S}_r(\pi) = \hat{S} = &\sum_{v=0}^{T} k(v/\ell(T)) \frac{1}{T}\Sigma_{v+1}^T m_{1t}(\hat{\theta}_1(1), \hat{\theta}_3(1), \hat{\tau}_1(1)) m_{1t-v}(\hat{\theta}_1(1), \hat{\theta}_3(1), \hat{\tau}_1(1))' \\ &+ \sum_{v=1}^{T} k(v/\ell(T)) \frac{1}{T}\Sigma_{v+1}^T m_{1t-v}(\hat{\theta}_1(1), \hat{\theta}_3(1), \hat{\tau}_1(1)) m_{1t}(\hat{\theta}_1(1), \hat{\theta}_3(1), \hat{\tau}_1(1))' \end{aligned}$$

for $r = 1, 2$, where $k(\cdot)$ is a kernel and $\ell(T)$ is a (possibly data–dependent) bandwidth parameter. See Andrews (1988) regarding the choice of kernel and bandwidth parameter. Corresponding to (2.16), kernel estimators $\hat{S}_1(\pi)$ and $\hat{S}_2(\pi)$ can be defined analogously using the data from the time periods $1, \ldots, T\pi$ and $T\pi+1, \ldots, T$, respectively, and using the estimators $(\hat{\theta}_1(\pi), \hat{\theta}_3(\pi), \hat{\tau}_1(\pi))$ and $(\hat{\theta}_2(\pi), \hat{\theta}_3(\pi), \hat{\tau}_2(\pi))$ respectively.

Under Assumption 2 and the following Assumption 3, the estimators $\hat{V}_r(\pi)$ are consistent uniformly over $\pi \in \Pi$:

ASSUMPTION 3: $\hat{V}_r(\pi)$ *is constructed using an estimator* $\hat{S}_r(\pi)$ *that satisfies* $\sup_{\pi\in\Pi}\|\hat{S}_r(\pi) - S\| \xrightarrow{p} 0$ *for* $r = 1, 2$.

THEOREM 3: *Under Assumptions* 2 *and* 3,

$$\sup_{\pi\in\Pi}\|\hat{V}_r(\pi) - V_r(\pi)\| \xrightarrow{p} 0 \quad \textit{for } r = 1, 2 ,$$

where $\hat{V}_r(\pi)$ *is as defined in* (2.15) *plus either* (2.16) *or* (2.17).

3. TESTS OF PARAMETER INSTABILITY

In this section, we introduce tests for parameter instability and structural change with unknown change point in general parametric models indexed by parameters $(\theta_{1t}, \theta_{30})$ for $t = 1, \ldots, T$. The null and alternative hypotheses of interest are

$$\text{(3.1)} \qquad \begin{aligned} &H_0 : \theta_{1t} = \theta_{10} \ \forall t \le T \ \text{ for some } \ \theta_{10} \in R^p . \\ &H_1 : \theta_{1s} \ne \theta_{1t} \ \text{ for some } \ s, t = 1, \ldots, T . \end{aligned}$$

Of particular interest are one–time structural change alternatives: For $\pi \in (0,1)$,

$$\text{(3.2)} \quad H_1(\pi) : \theta_{1t} = \begin{cases} \theta_1(\pi) & \text{for } t = 1,..,T\pi \\ \theta_2(\pi) & \text{for } t = T\pi+1,\ldots,T \end{cases} \quad \text{for some constants } \theta_1(\pi), \theta_2(\pi) \in R^p .$$

In the case of tests of pure structural change, no parameter θ_{30} appears and the entire parameter vector, viz. θ_{1t} , is subject to parameter instability under the alternative hypothesis H_1 or $H_1(\pi)$.

The first subsection below defines tests based on the Wald statistic. The second subsection defines tests based on type "a" LM–like and LR–like statistics. The third defines tests based on type "b" LM–like and LR–like statistics.

3.1. Tests Based on Wald Statistics

The Wald statistic for testing H_0 against $H_1(\pi)$ is given by

$$W_T(\pi) = T(\hat{\theta}_1(\pi) - \hat{\theta}_2(\pi))'(\hat{V}_1(\pi) + \hat{V}_2(\pi))^{-1}(\hat{\theta}_1(\pi) - \hat{\theta}_2(\pi)), \tag{3.3}$$

where $\hat{V}_1(\pi)$ and $\hat{V}_2(\pi)$ are as in (2.15) plus either (2.16) or (2.17) above.

Based on the Wald statistic $W_T(\pi)$, the following test statistic can be used for testing H_0 versus $\bigcup_{\pi\in\Pi} H_1(\pi)$ or H_0 versus H_1 :

$$\sup_{\pi\in\Pi} W_T(\pi), \tag{3.4}$$

where Π is a set with closure in (0,1). One rejects H_0 for large values of $\sup_{\pi\in\Pi} W_T(\pi)$. For reasons of power, the use of $\Pi = [0,1]$ is not desirable, see Comment 1 to Corollary 1 in Section 4 below. In addition, for $\Pi = [0,1]$ the asymptotic critical values given below are not valid, see Corollary 1.

The test statistic of (3.4) and those introduced below have the desired nuisance parameter–free asymptotic null distribution only under the following assumptions. Let a vector or matrix appended with a subscript $*$ denote the sub–matrix or sub-vector with blocks that correspond to θ_3 or $m_{3t}(\theta_{10}, \theta_{30}, \tau_0)$ deleted. Thus,

$$M_*(\pi) = \begin{bmatrix} \pi M & \underset{\sim}{0} \\ \underset{\sim}{0} & (1-\pi)M \end{bmatrix}, \quad D_*(\pi) = \begin{bmatrix} D_1(\pi) & D_{12}(\pi) \\ D_{21}(\pi) & D_2(\pi) \end{bmatrix}, \quad \text{and } x_* = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \tag{3.5}$$

where $x = (x_1', x_2', x_3')' \in R^{v_1} \times R^{v_1} \times R^{v_3} = R^v$, $M_*(\pi) \in R^{2v_1 \times 2p}$, and $D_*(\pi) \in R^{2v_1 \times 2v_1}$. Let

$$H = [I_p \vdots -I_p \vdots \underset{\sim}{0}] \in R^{p\times(2p\times p_3)} \text{ and } H_* = [I_p \vdots -I_p] \in R^{p\times 2p}. \tag{3.6}$$

ASSUMPITON 4: (i) S *is nonsingular and* M *is full rank* p $(\leq v_1)$.
(ii) $H(M(\pi)'D(\pi)M(\pi))^{-1}M(\pi)'D(\pi)x = H_*(M_*(\pi)'D_*(\pi)M_*(\pi))^{-1}M_*(\pi)'D_*(\pi)x_*$ $\forall x \in R^v$, $\forall \pi \in \Pi$.
(iii) $D_*(\pi)$ *is block diagonal with nonsingular* $v_1 \times v_1$ *blocks* $D_1(\pi)$ *and* $D_2(\pi)$ *that satisfy* $D_1(\pi) = a(\pi)D_2(\pi)$ *for some scalar constant* $a(\pi) \neq 0$.
(iv) *Either* $p = v_1$ (*i.e., the dimension of* θ_1 *equals the dimension of* $m_{1t}(\theta_1, \theta_3, \tau_1)$) *or* $D_r(\pi)$ *depends on* π *only through a scalar multiple for* $r = 1, 2$ (*i.e.,* $D_r(\pi) = \alpha_r(\pi)D_r$ *for some scalar* $\alpha_r(\pi)$ *and some matrix* D_r *for* $r = 1, 2$).

Three cases in which the key Assumption 4(ii) holds are as follows. (i) Assumption 4(ii) obviously holds in the context of tests for pure structural change, since $H = H_*$, $M(\pi) = M_*(\pi)$, etc. (ii) Assumption 4(ii) also holds in the case of "orthogonality" between the estimators of (θ_1, θ_2) and θ_3. That is, if $M(\pi)$ and $D(\pi)$ are block diagonal with blocks of dimension $(2v_1 \times 2p, v_3 \times p_3)$ and $(2v_1 \times 2v_1, v_3 \times v_3)$, respectively, then 4(ii) holds. (iii) If $M(\pi)$ is square (i.e., θ and $\bar{m}_T(\theta,\pi,\underset{\sim}{\tau})$ have the same dimension $2p + p_3 = v$) and nonsingular, then Assumption 4(ii) holds. In this case $(M(\pi)'D(\pi)M(\pi))^{-1}M(\pi)'D(\pi)x = M(\pi)^{-1}x$ and $M(\pi)$, defined in (2.10), is of the appropriate form for the application of Lemma A–3 of the Appendix.

EXAMPLE 1 (GMM, cont.): Suppose for the time being that a parameter θ_3 may appear in the GMM Example 1. In this scenario, Assumption 4(ii) does not always hold. It holds in the cases of (i) tests of pure structural change (i.e., when no θ_3 parameter appears), (ii) orthogonality between the estimators of (θ_1, θ_2) and θ_3, or (iii) equal dimensions of $(\theta_1', \theta_3')'$ and $f(W_t, \theta_1, \theta_3)$, i.e., $p + p_3 = v_1 + v_3$, and nonsingularity of $E\frac{\partial}{\partial(\theta_1', \theta_3')}f(W_t, \theta_{10}, \theta_{30})$ (where $f(W_t, \theta_1, \theta_3)$ is the v_1+v_3–vector function that defines the GMM estimator when a parameter θ_3 may be present).

We return now to the case where no parameter θ_3 appears in the GMM example. Assumption 4(i) holds by Assumption GMM–2(ii) and (vi). Also, Assumptions 4(iii) and

(iv) are satisfied since $D(\pi) = D_*(\pi)$ is block diagonal with blocks $D_1(\pi) = a_0' a_0/\pi$ and $D_2(\pi) = a_0' a_0/(1-\pi)$.

EXAMPLE 2 (ML, cont.): In this example, $M(\pi)$ is symmetric and nonsingular, so Assumption 4(ii) always holds. Assumption 4(i) holds because $M = S$ and M is nonsingular (since it is the upper $p \times p$ block of the nonsingular asymptotic information matrix for (θ_1, θ_3)). To obtain $M = S$, $f_t(\theta_1, \theta_3)$ must be sufficiently regular to permit interchange of the $\frac{\partial}{\partial(\theta_1', \theta_3')'}$ and $\int$ operations in the expressions $\frac{\partial}{\partial(\theta_1', \theta_3')'}\int\left[\frac{\partial}{\partial(\theta_1', \theta_3')}\log f_t(\theta_{10}, \theta_{30})\right]f_t(\theta_{10}, \theta_{30})d\xi_T$ and $\frac{\partial}{\partial(\theta_1', \theta_3')'}\int f_t(\theta_{10}, \theta_{30})d\xi_T$ (since in this case $\left\{\left[-\frac{\partial}{\partial(\theta_1', \theta_3')'}\log f_t(\theta_{10}, \theta_{30}), \sigma(\ldots, W_t)\right] : t \geq 1\right\}$ is a martingale difference sequence and the conditional information matrix equality holds). Assumptions 4(iii) and (iv) are satisfied since $D(\pi) = I_{2p+p_3}$ and $D_1(\pi) = D_2(\pi) = I_p$.

3.2. Tests Based on Type "a" LM–like and LR–like Statistics

In this subsection and the next, we consider type a and type b LM–like and LR–like test statistics. The LM_a and LR_a statistics are defined for any criterion function $d(\bar{m}_T(\theta, \pi, \hat{\tau}(\pi)), \hat{\gamma}(\pi))$ that satisfies Assumptions 2, 4, and 5a (stated below), although the LR_a statistic has the desired asymptotic null distribution only under special conditions. These conditions are sometimes satisfied in the GMM Example 1, but are not satisfied in the ML Example 2. The LM_b and LR_b statistics are defined only for criterion functions for which $d(m,\gamma) = m'm/2$ and $\bar{m}_T(\theta,\pi,\tau) = \frac{\partial}{\partial\theta}\bar{\rho}_T(\theta,\pi,\tau)$ for some function $\bar{\rho}_T(\theta,\pi,\tau)$. The latter conditions are satisfied in the ML Example 2 but not in the GMM Example 1. For ML estimators, the LM_b and LR_b statistics correspond to the usual LM and LR test statistics.

First we define the $LM_{aT}(\pi)$ and $LR_{aT}(\pi)$ statistics. They make use of the restricted extremum estimator $\tilde{\theta}_a(\cdot)$:

DEFINITION: A sequence of *restricted extremum estimators* $\{\tilde{\theta}_a(\cdot)\} = \{(\tilde{\theta}_a(\pi) : \pi \in \Pi) : T \geq 1\}$ is any sequence of stochastic processes such that

(3.7) $d(\bar{m}_T(\tilde{\theta}_a(\pi),\pi,\hat{\tau}(\pi)),\hat{\gamma}(\pi)) = \inf\{d(\bar{m}_T(\theta,\pi,\hat{\tau}(\pi)),\hat{\gamma}(\pi)) : \theta = (\theta_1',\theta_2',\theta_3')' \in \Theta,\ \theta_1 = \theta_2\}$

with probability $\to 1$.

Note that the same preliminary estimators $\hat{\tau}(\pi)$ and $\hat{\gamma}(\pi)$ are used to define $\tilde{\theta}_a(\pi)$ as are used to define $\hat{\theta}(\pi)$. This is necessary for the $LM_{aT}(\pi)$ and $LR_{aT}(\pi)$ test statistics to have χ^2 asymptotic null distributions for each fixed π.

Suppose the null hypothesis H_0 is true. If Assumption 1 holds for the parameter space Θ, it also holds for the parameter space $\tilde{\Theta}_0 = \{\theta \in \Theta : \theta_1 = \theta_2\}$. Thus, Assumption 1 and Theorem 1 imply that $\sup_{\pi\in\Pi}\|\tilde{\theta}_a(\pi) - \theta_0\| \xrightarrow{p} 0$. In consequence, the first part of the following assumption is straightforward to verify:

ASSUMPTION 5a: (i) *If* θ_0 *satisfies the null hypothesis,* $\sup_{\pi\in\Pi}\|\tilde{\theta}_a(\pi) - \theta_0\| \xrightarrow{p} 0$.
(ii) $H(M(\pi)'D(\pi)M(\pi))^{-1}H' = H_*(M_*(\pi)'D_*(\pi)M_*(\pi))^{-1}H_*' \quad \forall \pi \in \Pi$.

The second part of Assumption 5a is closely related to Assumption 4(ii). It holds in the case of testing for pure structural change. It also holds in the case of orthogonality between the estimators of (θ_1, θ_2) and θ_3 (as defined above). In addition, it holds if $M(\pi)$ is square $(2p + p_3 = v)$ and nonsingular and $D(\pi)$ is block diagonal with blocks of dimension $(2v_1 \times 2v_1, v_3 \times v_3)$.[7] Thus, in the GMM Example 1, Assumption 5a(ii) is satisfied in some, but not all, cases.

The LM_a statistic uses estimators of $V_1(\pi)$ and $V_2(\pi)$ that are constructed with the restricted estimator $\tilde{\theta}_a(\cdot)$ in place of $\hat{\theta}(\cdot)$. Let $\tilde{V}_r(\pi)$, $\tilde{M}_r(\pi)$, and $\tilde{D}_r(\pi)$ be defined as are $\hat{V}_r(\pi)$, $\hat{M}_r(\pi)$, and $\hat{D}_r(\pi)$ in (2.15) plus either (2.16) or (2.17) but with $\hat{\theta}(\cdot)$ replaced by $\tilde{\theta}_a(\cdot)$.

For fixed change point π, the LM_a statistic is defined to be

$$
\begin{aligned}
(3.8)\qquad \mathrm{LM}_{\mathrm{aT}}(\pi) = \mathrm{T}\Big[&\tfrac{\partial}{\partial\theta_1'}\mathrm{d}(\bar{\mathrm{m}}_{\mathrm{T}}(\tilde\theta_{\mathrm{a}}, \pi, \hat{\tau}), \hat\gamma)(\tilde{\mathrm{M}}_1'\tilde{\mathrm{D}}_1\tilde{\mathrm{M}}_1)^{-1}\tilde\delta_1^{-2} \\
&- \tfrac{\partial}{\partial\theta_2'}\mathrm{d}(\bar{\mathrm{m}}_{\mathrm{T}}(\tilde\theta_{\mathrm{a}}, \pi, \hat{\tau}), \hat\gamma)(\tilde{\mathrm{M}}_2'\tilde{\mathrm{D}}_2\tilde{\mathrm{M}}_2)^{-1}\tilde\delta_2^{-2}\Big] \\
&\times(\tilde{\mathrm{V}}_1 + \tilde{\mathrm{V}}_2)^{-1}\Big[\tilde\delta_1^{-2}(\tilde{\mathrm{M}}_1'\tilde{\mathrm{D}}_1\tilde{\mathrm{M}}_1)^{-1}\tfrac{\partial}{\partial\theta_1}\mathrm{d}(\bar{\mathrm{m}}_{\mathrm{T}}(\tilde\theta_{\mathrm{a}}, \pi, \hat{\tau}), \hat\gamma) \\
&- \tilde\delta_2^{-2}(\tilde{\mathrm{M}}_2'\tilde{\mathrm{D}}_2\tilde{\mathrm{M}}_2)^{-1}\tfrac{\partial}{\partial\theta_2}\mathrm{d}(\bar{\mathrm{m}}_{\mathrm{T}}(\tilde\theta_{\mathrm{a}}, \pi, \hat{\tau}), \hat\gamma)\Big],
\end{aligned}
$$

where here and below the dependence on π of $\tilde\theta_{\mathrm{a}}$, $\hat{\tau}$, $\hat\gamma$, $\tilde{\mathrm{M}}_{\mathrm{r}}$, $\tilde{\mathrm{D}}_{\mathrm{r}}$, $\tilde\delta_{\mathrm{r}}$, and $\tilde{\mathrm{V}}_{\mathrm{r}}$ is suppressed for notational simplicity.

Next, for fixed change point π, the LR_{a} statistic is defined by

$$(3.9)\qquad \mathrm{LR}_{\mathrm{aT}}(\pi) = 2\mathrm{T}(\mathrm{d}(\bar{\mathrm{m}}_{\mathrm{T}}(\tilde\theta_{\mathrm{a}}(\pi), \pi, \hat{\tau}), \hat\gamma) - \mathrm{d}(\bar{\mathrm{m}}_{\mathrm{T}}(\hat\theta(\pi), \pi, \hat{\tau}), \hat\gamma))/\hat{\mathrm{b}},$$

where $\hat{\mathrm{b}}$ is a scalar rv defined in Assumption 6a below. The preliminary estimators $(\hat{\tau}(\pi), \hat\gamma(\pi))$ used in $\mathrm{LR}_{\mathrm{aT}}(\pi)$ may be restricted or unrestricted estimators of $(\tau_0, \gamma_0(\pi))$. They must be the same in both criterion functions used to calculate $\mathrm{LR}_{\mathrm{aT}}(\pi)$, however, and they must be such that both $\hat\theta(\pi)$ and $\tilde\theta_{\mathrm{a}}(\pi)$ are consistent under the null hypothesis.

The $\mathrm{LR}_{\mathrm{aT}}(\pi)$ statistic has the desired asymptotic χ^2 null distribution for each fixed π only under the following assumption:

ASSUMPTION 6a: $\mathrm{M}'\mathrm{D}_{\mathrm{r}}(\pi)\mathrm{S}\mathrm{D}_{\mathrm{r}}(\pi)\mathrm{M} = \mathrm{b}\delta_{\mathrm{r}}^{-1}(\pi)\mathrm{M}'\mathrm{D}_{\mathrm{r}}(\pi)\mathrm{M}$ *for* $\mathrm{r} = 1, 2$, *for some scalar constant* $\mathrm{b} \neq 0$ *and* $\hat{\mathrm{b}} \xrightarrow{p} \mathrm{b}$ *for some sequence of non-zero rv's* $\{\hat{\mathrm{b}}\}$.

When Assumption 6a holds, $\hat{\mathrm{V}}_{\mathrm{r}}(\pi)$ can be simplified to

$$(3.10)\qquad \hat{\mathrm{V}}_{\mathrm{r}}(\pi) = \hat{\mathrm{b}}\hat\delta_{\mathrm{r}}^{-2}(\pi)(\hat{\mathrm{M}}_{\mathrm{r}}(\pi)'\hat{\mathrm{D}}_{\mathrm{r}}(\pi)\hat{\mathrm{M}}_{\mathrm{r}}(\pi))^{-1} \text{ for } \mathrm{r} = 1, 2.$$

$\tilde{\mathrm{V}}_{\mathrm{r}}(\pi)$ can be simplified analogously. In this case, $\mathrm{LM}_{\mathrm{aT}}(\pi)$ simplifies to

$$
\begin{aligned}
(3.11)\qquad \mathrm{LM}_{\mathrm{aT}}(\pi) \doteq \sum_{\mathrm{r}=1}^{2} &\mathrm{T}\tfrac{\partial}{\partial\theta_{\mathrm{r}}'}\mathrm{d}(\bar{\mathrm{m}}_{\mathrm{T}}(\tilde\theta_{\mathrm{a}}(\pi), \pi, \hat{\tau}), \hat\gamma)\tilde\delta_{\mathrm{r}}^{-2}(\pi)(\tilde{\mathrm{M}}_{\mathrm{r}}(\pi)'\tilde{\mathrm{D}}_{\mathrm{r}}(\pi)\tilde{\mathrm{M}}_{\mathrm{r}}'(\pi))^{-1} \\
&\times\tfrac{\partial}{\partial\theta_{\mathrm{r}}}\mathrm{d}(\bar{\mathrm{m}}_{\mathrm{T}}(\tilde\theta_{\mathrm{a}}(\pi), \pi, \hat{\tau}), \hat\gamma)/\hat{\mathrm{b}},
\end{aligned}
$$

where $\doteq$ denotes equality that holds with probability $\to 1$. (This simplification of $LM_{aT}(\pi)$ occurs because $\frac{\partial}{\partial\theta}d(\bar{m}_T(\tilde{\theta}_a(\pi), \pi, \hat{\tau}), \hat{\gamma}) \doteq -[I_p \vdots -I_p \vdots 0]'\tilde{\lambda}(\pi)$ for some p–vector $\tilde{\lambda}(\pi)$ of Lagrange multipliers.)

As in (3.4), for testing H_0 versus $\bigcup_{\pi\in\Pi} H_1(\pi)$ or H_0 versus H_1, we consider the statistics

$$\sup_{\pi\in\Pi} LM_{aT}(\pi) \text{ and } \sup_{\pi\in\Pi} LR_{aT}(\pi)\,. \tag{3.12}$$

The null hypothesis H_0 is rejected for large values of these statistics.

EXAMPLE 1 (GMM, cont.): Here, the major components of $LM_{aT}(\pi)$ are given by

$$\begin{aligned}&\frac{\partial}{\partial\theta_1}d(\bar{m}_T(\tilde{\theta}_a(\pi),\pi,\hat{\tau}), \hat{\gamma}) = \tilde{M}_1(\pi)'\frac{1}{\pi^2}a_{1,T\pi}'a_{1,T\pi}\frac{1}{T}\Sigma_1^{T\pi}f(W_t, \tilde{\theta}_{1a}(\pi)) \text{ and}\\ &\frac{\partial}{\partial\theta_2}d(\bar{m}_T(\tilde{\theta}_a(\pi),\pi,\hat{\tau}), \hat{\gamma}) = \tilde{M}_2(\pi)'\frac{1}{(1-\pi)^2}a_{2,T\pi}'a_{2,T\pi}\frac{1}{T}\Sigma_{T\pi+1}^{T}f(W_t, \tilde{\theta}_{1a}(\pi))\,,\end{aligned} \tag{3.13}$$

where $\tilde{\theta}_a(\pi) = (\tilde{\theta}_{1a}(\pi)', \tilde{\theta}_{1a}(\pi)')'$.

In this example, Assumption 6a holds with $b = 1$ if

$$a_0'a_0 = S^{-1}\,. \tag{3.14}$$

In this case, $\hat{V}_r(\pi)$, $\tilde{V}_r(\pi)$, and $LM_{aT}(\pi)$ can be simplified as in (3.10) and (3.11), respectively, and $LR_{aT}(\pi)$ is given by

$$\begin{aligned}LR_{aT}(\pi) = T\Big[&\frac{1}{T}\Sigma_1^{T\pi}f(W_t, \tilde{\theta}_{1a}(\pi))'\frac{1}{\pi}a_{1,T\pi}'a_{1,T\pi}\frac{1}{T}\Sigma_1^{T\pi}f(W_t, \tilde{\theta}_{1a}(\pi))\\ &+\frac{1}{T}\Sigma_{T\pi+1}^{T}f(W_t, \tilde{\theta}_{1a}(\pi))'\frac{1}{1-\pi}a_{2,T\pi}'a_{2,T\pi}\frac{1}{T}\Sigma_{T\pi+1}^{T}f(W_t, \tilde{\theta}_{1a}(\pi))\\ &-\frac{1}{T}\Sigma_1^{T\pi}f(W_t, \hat{\theta}_1(\pi))'\frac{1}{\pi}a_{1,T\pi}'a_{1,T\pi}\frac{1}{T}\Sigma_1^{T\pi}f(W_t, \hat{\theta}_1(\pi))\\ &-\frac{1}{T}\Sigma_{T\pi+1}^{T}f(W_t, \hat{\theta}_2(\pi))'\frac{1}{1-\pi}a_{2,T\pi}'a_{2,T\pi}\frac{1}{T}\Sigma_{T\pi+1}^{T}f(W_t, \hat{\theta}_2(\pi))\Big]\,.\end{aligned} \tag{3.15}$$

3.2. Tests Based on Type b LM–like and LR–like Statistics

First, we describe the context in which the LM_b and LR_b statistics are defined:

ASSUMPTION 6b: (i) $d(m,\gamma) = m'm/2$. *There exist functions* $\{\rho_{Tt}(W_{Tt}, \theta_1, \theta_3, \tau_1)\}$ *such that* $\frac{\partial}{\partial\theta_r}\rho_{Tt}(W_{Tt}, \theta_1, \theta_3, \tau_1) = m_{rTt}(W_{Tt}, \theta_1, \theta_3, \tau_1)$ $\forall t$, *for* $r = 1, 3$. *With probability* $\to 1$, $\hat{\theta}(\pi)$ *solves* $\inf\{\bar{\rho}_T(\theta, \pi, \hat{\tau}(\pi)) : \theta \in \Theta\}$, *where* $\bar{\rho}_T(\theta,\pi,\tau) = \frac{1}{T}\Sigma_1^{T\pi}\rho_{Tt}(W_{Tt}, \theta_1, \theta_3, \tau_1) + \frac{1}{T}\Sigma_{T\pi+1}^{T}\rho_{Tt}(W_{Tt}, \theta_2, \theta_3, \tau_2)$.
(ii) $S = cM$ *for some scalar* $c \neq 0$ *and* $\hat{c} \xrightarrow{p} c$ *for some sequence of non–zero rv's* $\{\hat{c}\}$.

When Assumption 6b(i) holds the $LM_{bT}(\pi)$ and $LR_{bT}(\pi)$ statistics are defined. The $LR_{bT}(\pi)$ statistic has an asymptotic χ^2 null distribution for fixed π, however, only when Assumption 6b(ii) also holds.

EXAMPLE 2 (ML, cont.): In the ML case, Assumption 6b(i) holds with

$$\rho(W_{Tt}, \theta_1, \theta_3, \tau_1) = -\log f_t(\theta_1, \theta_3). \tag{3.16}$$

Assumption 6b(ii) holds with $c = 1$, since $S = M$ (see the end of Section 3.1).

The $LM_{bT}(\pi)$ and $LR_{bT}(\pi)$ statistics make use of the restricted estimator $\tilde{\theta}_b(\pi)$:

DEFINITION: A sequence of *restricted extremum estimators* $\{\tilde{\theta}_b(\cdot)\} = \{(\tilde{\theta}_b(\pi) : \pi \in \Pi) : T \geq 1\}$ is any sequence of stochastic processes such that

$$\bar{\rho}_T(\tilde{\theta}_b(\pi), \pi, \hat{\tau}(\pi)) = \inf\{\bar{\rho}_T(\theta, \pi, \hat{\tau}(\pi)) : \theta = (\theta_1', \theta_2', \theta_3')' \in \Theta, \theta_1 = \theta_2\} \tag{3.17}$$

with probability $\to 1$.

Note that if $\hat{\tau}(\pi)$ does not depend on π or if $\rho_{Tt}(\cdot,\cdot,\cdot)$ does not depend on a preliminary nuisance parameter estimator $\hat{\tau}(\pi)$, then $\tilde{\theta}_b(\pi)$ does not depend on π. Also note that $\tilde{\theta}_a(\pi)$ and $\tilde{\theta}_b(\pi)$ differ in general, because $\tilde{\theta}_b(\pi)$ minimizes $\bar{\rho}_T(\theta,\pi,\hat{\tau}(\pi))$ subject to $\theta_1 = \theta_2$, whereas $\tilde{\theta}_a(\pi)$ minimizes $\frac{\partial}{\partial\theta'}\bar{\rho}_T(\theta,\pi,\hat{\tau}(\pi))\frac{\partial}{\partial\theta}\bar{\rho}_T(\theta,\pi,\hat{\tau}(\pi))$ subject to $\theta_1 = \theta_2$.

As with $\{\tilde{\theta}_a(\pi)\}$, the consistency of $\{\tilde{\theta}_b(\pi)\}$ can be established using Theorem 1. Thus, the following assumption is straightforward to verify:

ASSUMPTION 5b: *If* θ_0 *satisfies the null hypothesis,* $\sup_{\pi\in\Pi}\|\tilde{\theta}_b(\pi)-\theta_0\| \overset{p}{\to} 0$.

By definition,

$$\text{(3.18)}\qquad \begin{aligned} LM_{bT}(\pi) = T\Big[&\tfrac{\partial}{\partial\theta_1'}\bar{\rho}_T(\tilde{\theta}_b,\pi,\hat{\underset{\sim}{\tau}})\delta_1^{-1}\tilde{M}_1^{-1} - \tfrac{\partial}{\partial\theta_2'}\bar{\rho}_T(\tilde{\theta}_b,\pi,\hat{\underset{\sim}{\tau}})\delta_2^{-1}\tilde{M}_2^{-1}\Big] \\ &\times\Big[\delta_1^{-1}\tilde{M}_1^{-1}\hat{S}_1\tilde{M}_1^{-1} + \delta_2^{-1}\tilde{M}_2^{-1}\hat{S}_2\tilde{M}_2^{-1}\Big]^{-1} \\ &\times\Big[\delta_1^{-1}\tilde{M}_1^{-1}\tfrac{\partial}{\partial\theta_1'}\bar{\rho}_T(\tilde{\theta}_b,\pi,\hat{\underset{\sim}{\tau}}) - \delta_2^{-1}\tilde{M}_2^{-1}\tfrac{\partial}{\partial\theta_2'}\bar{\rho}_T(\tilde{\theta}_b,\pi,\hat{\underset{\sim}{\tau}})\Big] \quad\text{and} \\ LR_{bT}(\pi) = 2T(&\bar{\rho}_T(\tilde{\theta}_b,\pi,\hat{\underset{\sim}{\tau}}) - \bar{\rho}_T(\hat{\theta},\pi,\hat{\underset{\sim}{\tau}}))/\hat{c}\,, \end{aligned}$$

where $\tilde{M}_r$ $(=\tilde{M}_r(\pi))$ and $\hat{S}_r$ $(=\hat{S}_r(\pi))$ are as defined above but with $\tilde{\theta}_b(\pi)$ in place of $\tilde{\theta}_a(\pi)$ or $\hat{\theta}(\pi)$. To obtain χ^2 asymptotic null distributions for fixed π, the preliminary estimator $\hat{\underset{\sim}{\tau}}(\pi)$ must be the same in both criterion functions used to calculate $LR_{bT}(\pi)$ and used to define $\tilde{\theta}_b(\pi)$ and $\hat{\theta}(\pi)$.

When Assumptions 2 and 6b(i) hold, we have: $D_r(\pi)=I_p$ for $r=1,2$, $M=\lim_{T\to\infty}\frac{1}{T}\Sigma_1^T E\frac{\partial^2}{\partial\theta_1\partial\theta_1'}\rho_{Tt}(W_{Tt},\theta_{10},\theta_{30},\tau_0)$, and M is nonsingular by Assumption 2(h). If in addition $\hat{S}_r(\pi)=\hat{c}\tilde{M}_r(\pi)$ $\forall\pi\in\Pi$, for $r=1,2$, for some scalar rv $\hat{c}\neq 0$, as usually occurs when Assumption 6b(ii) holds, then $LM_{bT}(\pi)$ simplifies as follows:

$$\text{(3.19)}\qquad LM_{bT}(\pi) \doteq \sum_{r=1}^{2} T\frac{\partial}{\partial\theta_r'}\bar{\rho}_T(\tilde{\theta}_b(\pi),\pi,\hat{\underset{\sim}{\tau}})\delta_r^{-1}\tilde{M}_r(\pi)^{-1}\frac{\partial}{\partial\theta_r}\bar{\rho}_T(\tilde{\theta}_b(\pi),\pi,\hat{\underset{\sim}{\tau}})/\hat{c}\,.$$

(This simplification uses the fact that $\frac{\partial}{\partial\theta}\bar{\rho}_T(\tilde{\theta}_b(\pi),\pi,\hat{\underset{\sim}{\tau}}) \doteq -[I_p \vdots -I_p \vdots \underset{\sim}{0}]'\tilde{\eta}(\pi)$ for some p–vector $\tilde{\eta}(\pi)$ of Lagrange multipliers.)

For testing H_0 versus $\cup_{\pi\in\Pi} H_1(\pi)$ or H_0 versus H_1, we consider the statistics

$$\text{(3.20)}\qquad \sup_{\pi\in\Pi} LM_{bT}(\pi) \quad\text{and}\quad \sup_{\pi\in\Pi} LR_{bT}(\pi)\,.$$

The null hypothesis H_0 is rejected for large values of these statistics.

EXAMPLE 2 (ML, cont.): In the ML case, we have

$$\mathrm{LM}_{bT}(\pi) = T\Big[\tfrac{1}{T}\Sigma_1^{T\pi} \tfrac{\partial}{\partial\theta_1}\log f_t(\tilde\theta_{1b}, \tilde\theta_{3b})\Big]' \Big[-\tfrac{1}{T}\Sigma_1^{T\pi} \tfrac{\partial^2}{\partial\theta_1\partial\theta_1'}\log f_t(\tilde\theta_{1b}, \tilde\theta_{3b})\Big]^{-1}$$
$$\times\tfrac{1}{T}\Sigma_1^{T\pi} \tfrac{\partial}{\partial\theta_1}\log f_t(\tilde\theta_{1b}, \tilde\theta_{3b}) + T\Big[\tfrac{1}{T}\Sigma_{T\pi+1}^{T} \tfrac{\partial}{\partial\theta_1}\log f_t(\tilde\theta_{1b}, \tilde\theta_{3b})\Big]'$$
$$\times\Big[-\tfrac{1}{T-T\pi}\Sigma_{T\pi+1}^{T} \tfrac{\partial^2}{\partial\theta_1\partial\theta_1'}\log f_t(\tilde\theta_{1b}, \tilde\theta_{3b})\Big]^{-1}\tfrac{1}{T}\Sigma_1^{T\pi} \tfrac{\partial}{\partial\theta_1}\log f_t(\tilde\theta_{1b}, \tilde\theta_{3b}), \tag{3.21}$$

$$\mathrm{LR}_{bT}(\pi) = -2\Big[\Sigma_1^{T}\log f_t(\tilde\theta_{1b}, \tilde\theta_{3b}) - \Sigma_1^{T\pi}\log f_t(\hat\theta_1(\pi), \hat\theta_3(\pi))$$
$$-\Sigma_{T\pi+1}^{T}\log f_t(\hat\theta_2(\pi), \hat\theta_3(\pi))\Big],$$

where $\tilde\theta_b = (\tilde\theta_{1b}', \tilde\theta_{1b}', \tilde\theta_{3b}')'$ does not depend on π since no estimator $\hat{\underset{\sim}{\gamma}}(\pi)$ appears.

4. ASYMPTOTIC PROPERTIES OF THE TEST STATISTICS

4.1. Asymptotic Distributions under the Null Hypothesis

The first result of this section provides the asymptotic null distributions of the test statistics introduced in Section 3.

THEOREM 4: *Suppose Assumptions* 2–4 *hold. Given any set* Π *whose closure lies in* (0,1), *the following processes indexed by* $\pi \in \Pi$ *satisfy:*

(a) $W_T(\cdot) \Rightarrow Q_p(\cdot)$ *and* $\sup_{\pi\in\Pi} W_T(\pi) \xrightarrow{d} \sup_{\pi\in\Pi} Q_p(\pi)$, *where* $Q_p(\cdot) = (B^*(\cdot) - \iota(\cdot)B^*(1))'(B^*(\cdot) - \iota(\cdot)B^*(1))/[\iota(\cdot)(1-\iota(\cdot))]$,

(b) $\mathrm{LM}_{aT}(\cdot) \Rightarrow Q_p(\cdot)$ *and* $\sup_{\pi\in\Pi} \mathrm{LM}_{aT}(\pi) \xrightarrow{d} \sup_{\pi\in\Pi} Q_p(\pi)$ *provided Assumption* 5a *also holds,*

(c) $\mathrm{LR}_{aT}(\cdot) \Rightarrow Q_p(\cdot)$ *and* $\sup_{\pi\in\Pi} \mathrm{LR}_{aT}(\pi) \xrightarrow{d} \sup_{\pi\in\Pi} Q_p(\pi)$ *provided Assumptions* 5a *and* 6a *hold in place of* 3,

(d) $\mathrm{LM}_{bT}(\cdot) \Rightarrow Q_p(\cdot)$ *and* $\sup_{\pi\in\Pi} \mathrm{LM}_{bT}(\pi) \xrightarrow{d} \sup_{\pi\in\Pi} Q_p(\pi)$ *provided Assumptions* 5b *and* 6b(i) *also hold, and*

(e) $LR_{bT}(\cdot) \Rightarrow Q_p(\cdot)$ *and* $\sup_{\pi \in \Pi} LR_{bT}(\pi) \xrightarrow{d} \sup_{\pi \in \Pi} Q_p(\pi)$ *provided Assumptions* 5b *and* 6b *hold in place of* 3,

where $B^*(\cdot)$ *is a* p*–vector of independent Brownian motions on* [0,1] *restricted to* Π *and* $\iota(\cdot)$ *is the identity function, i.e.,* $\iota(\pi) = \pi \;\; \forall \pi \in \Pi$. *Furthermore, the convergence in* (a)–(e) *holds jointly.*

COMMENTS: 1. The limit process $Q_p(\cdot)$ is referred to in the literature as the square of a standardized tied–down Bessel process of order p, see Sen (1981, p. 46). For any fixed $\pi \in (0,1)$, $Q_p(\pi)$ has a chi–square distribution with p degrees of freedom. Under the assumptions, the asymptotic null distributions of $\sup_{\pi \in \Pi} W_T(\pi), \ldots, \sup_{\pi \in \Pi} LR_{bT}(\pi)$ are free of nuisance parameters except for the dimension p of θ_1. Thus, critical values for the test statistics can be tabulated, see below.

2. If $\Pi = [\pi_1, \pi_2]$ for $0 < \pi_1 \leq \pi_2 < 1$, then it can be shown (e.g., see the proof of Corollary 1 below) that

$$\text{(4.1)} \qquad P\left[\sup_{\pi \in \Pi} Q_p(\pi) > c_\alpha\right] = P\left[\sup_{s \in [1, \pi_2(1-\pi_1)/(\pi_1(1-\pi_2))]} BM(s)'BM(s)/s > c_\alpha\right],$$

where $BM(\cdot)$ denotes a p–vector of independent Brownian motion processes on $[0,\infty)$. In consequence, critical values based on the distribution of $\sup_{\pi \in [\pi_1, \pi_2]} Q_p(\pi)$ depend on π_1 and π_2 only through the parameter $\lambda = \pi_2(1-\pi_1)/(\pi_1(1-\pi_2))$. This simplifies the tabulation of such critical values.

Since $\|BM(\cdot)\|$ is a Bessel process of order p, the probability on the right–hand side of (4.1) is the probability that a Bessel process exceeds a square root boundary somewhere in the given interval. Such probabilities and corresponding critical values for given significance levels have been computed numerically for $p \leq 4$ by DeLong (1981).

3. Critical values c_α for the test statistics $\sup_{\pi \in \Pi} W_T(\pi), \ldots, \sup_{\pi \in \Pi} LR_{bT}(\pi)$ are provided in Table 1 based on their asymptotic null distribution $\sup_{\pi \in \Pi} Q_p(\pi)$. By

definition, c_α satisfies $P(\sup_{\pi\in\Pi} Q_p(\pi) > c_\alpha) = \alpha$. The tables cover significance levels $\alpha = .01, .025, .05$, and $.10$, $p = 1(1)20$, and $\Pi = [.15, .85]$. The given choice of Π was determined (subjectively) by trading off the length of Π and the power of the test for one-time structural change for points in Π in the Monte Carlo experiment described in Section 5.

The values reported in Table 1 are estimates of the critical values c_α obtained by (i) approximating the distribution of the supremum of $Q_p(\pi)$ over $\pi \in [.15, .85]$ by its maximum over a fine grid of points $\Pi(N)$ and (ii) simulating the distribution of $\max_{\pi\in\Pi(N)} Q_p(\pi)$ by Monte Carlo. The grid $\Pi(N)$ is defined by

$$\Pi(N) = [.15, .85] \cap \{\pi = j/N : j = 0, 1, \ldots, N\}. \tag{4.2}$$

The value of N was chosen to be 3,600 based on a comparison of the approximations obtained here with the numerical results of DeLong (1981), which are available for $p \leq 4$. A single realization from the distribution of $\max_{\pi\in\Pi(N)} Q_p(\pi)$ was obtained by simulating a p–vector $B^*(\cdot)$ of independent Brownian motions at the discrete points in $\Pi(N)$ and computing $\max_{\pi\in\Pi(N)} (B^*(\pi) - \pi B^*(1))'(B^*(\pi) - \pi B^*(1))/[\pi(1-\pi)]$. The number of repetitions R used was 25,000. The accuracy of the simulated critical value for approximating the critical value based on the statistic $\max_{\pi\in\Pi(N)} Q_p(\pi)$ can be determined by noting that the rejection probability of the statistic $\max_{\pi\in\Pi(N)} Q_p(\pi)$ using the simulated critical value has mean α and standard error approximately equal to $(\alpha(1-\alpha)/R)^{1/2}$. For $\alpha = .01$, $.025$, $.05$, and $.10$, the standard errors are .001, .002, .002, and .003 respectively.

4. The requirement that Π is bounded away from zero and one is made to ensure that the functions mapping $\nu_T(\cdot)$ of Assumption 2(e) into each of the processes in Theorem 4(a)–(e) are continuous. For example, if $\Pi = [0,1]$, the functions $\pi \to 1/\pi$ and $\pi \to 1/(1-\pi)$ are not continuous. In fact, if $\Pi = [0,1]$, the test statistics

$\sup_{\pi\in\Pi} W_T(\pi), \ldots, \sup LR_{bT}(\pi)$ do not converge in distribution, see Corollary 1 below.

5. The p–dimensional Brownian motion process $B^*(\cdot)$ differs from the v_1–dimensional Brownian motion process $B_1(\cdot)$ of Theorem 2, unless $p = v_1$. If one considers joint weak convergence of the processes in Theorems 2 and 4, then $B^*(\cdot)$ has to be a matrix multiple of $B_1(\cdot)$ when $p < v_1$, see the proof of Theorem 4.

6. Suppose the alternatives of greatest interest are ones in which some sub–vector of θ_1, say $\theta_{1\ell}$, exhibits parameter instability. In this case, one can redefine θ_1 and θ_3 such that the new vector θ_1 only contains $\theta_{1\ell}$ and use the W, LM, or LR statistic to test for instability of $\theta_{1\ell}$. Alternatively, one can leave θ_1 and θ_3 defined as they are and base a Wald test on $\hat\theta_{1\ell}(\pi) - \hat\theta_{2\ell}(\pi)$, where $\hat\theta_{r\ell}(\pi)$ is the appropriate subvector of $\hat\theta_r(\pi)$ for $r = 1, 2$. In particular, let

$$(4.3)\qquad W_{\ell T}(\pi) = T(\hat\theta_{1\ell}(\pi) - \hat\theta_{2\ell}(\pi))'([\hat V_1(\pi)]_{\ell\ell} + [\hat V_2(\pi)]_{\ell\ell})^{-1}(\hat\theta_{1\ell}(\pi) - \hat\theta_{2\ell}(\pi)),$$

where $[\hat V_r(\pi)]_{\ell\ell}$ denotes the sub–matrix of $\hat V_r(\pi)$ that corresponds to $\hat\theta_{r\ell}(\pi)$ for $r = 1, 2$. Under H_0 and Assumptions 2–4, $W_{\ell T}(\cdot)$ and $\sup_{\pi\in\Pi} W_{\ell T}(\pi)$ each have the same asymptotic distribution as that given for $W_T(\cdot)$ and $\sup_{\pi\in\Pi} W_T(\pi)$ in Theorem 4(a) but with p replaced by the dimension of $\theta_{1\ell}$. Note that the null hypothesis H_0 must be as defined in (3.1) for the above result to hold. The null hypothesis cannot include cases where $\theta_{1\ell}$ is constant over the sample, but other coefficients are variable.

7. If $m_{rt}(\theta_1, \theta_3, \tau_1)$ is not differentiable in θ_1, θ_3, or τ_1 for $r = 1, 3$ (Assumption 2(g)), e.g., as occurs with least absolute deviation estimators, then analogous results to those of Theorem 4 can be obtained for the $W_T(\cdot)$ statistic using the stochastic equicontinuity approach of Andrews (1989a,b,c), provided $Em_{rt}(\theta_1, \theta_3, \tau_1)$ is differentiable in θ_1, θ_3, and τ_1. Alternatively, if τ_1 is infinite dimensional, and hence, $\hat\theta$ is a semiparametric estimator, then analogous results to those of Theorem 4 can be obtained for

all of the test statistics $W_T(\cdot), \ldots, LR_{bT}(\cdot)$ using the approach of Andrews (1989a,b,c).

8. The fluctuation test of Sen (1980) and Ploberger, Krämer, and Kontrus (1989) can be extended to general parametric models as follows. Let

$$\text{(4.4)} \qquad FL_T(\pi) = \max_{j \leq p} \pi |[\hat{V}_1(1)^{-1/2}\sqrt{T}(\hat{\theta}_1(\pi) - \hat{\theta}_1(1))]_j| \ ,$$

where $\hat{V}_1(1)$ is as defined in (2.15) and $[\underset{\sim}{c}]_j$ denotes the j–th element of the p–vector $\underset{\sim}{c}$. Analogously to the results of Theorem 4, the fluctuation test statistic $\sup_{\pi \in \Pi} FL_T(\pi)$ satisfies

$$\text{(4.5)} \qquad \sup_{\pi \in \Pi} FL_T(\pi) \xrightarrow{d} \sup_{\pi \in \Pi} \sup_{j \leq p} |B_j^*(\pi) - \pi B_j^*(1)| \ ,$$

where $B^*(\pi) = (B_1^*(\pi), \ \ldots, \ B_p^*(\pi))'$, under the null hypothesis and under the assumptions given in Andrews (1989e).

9. The results of Theorem 4 also establish the asymptotic distributions of test statistics of the form $g(\{W_T(\pi) : \pi \in \Pi\})$ for arbitrary continuous functions g (using the uniform metric on the space of real functions on Π). In particular, $g(\{W_T(\pi) : \pi \in \Pi\}) \Rightarrow g(\{Q_p(\pi) : \pi \in \Pi\})$ under the assumptions and likewise for $LM_{aT}(\cdot)$ etc.

EXAMPLE 1 (GMM, cont.): Utilizing previous results, we now provide a set of sufficient conditions for the assumptions of Theorem 4 to hold in the GMM context. First, Assumptions GMM–1 and GMM–2 imply Assumptions 2, 4, and 5a. Next, $a_0'a_0 = S^{-1}$ implies Assumption 6a. Lastly, if $\hat{S}_r(\pi) = (a_{r,T\pi}'a_{r,T\pi})^{-1}$ for $r = 1, 2$, then Assumption 3 holds.

EXAMPLE 2 (ML, cont.): For the ML case, the conditions given at the end of Sections 2.2 and 2.3 are sufficient for Assumptions 2, 4, and 5b. These conditions plus that given at the end of Section 3.1 are sufficient for Assumption 6b. Lastly, Assumption 3 holds when Assumption 6b holds, if $\hat{S}_r(\pi) = \hat{M}_r(\pi)$ or $\hat{S}_r(\pi) = \tilde{M}_r(\pi)$ for $r = 1, 2$.

Next, we consider the limiting behavior of the statistics $\sup_{\pi\in\Pi} W_T(\pi)$, ..., $\sup_{\pi\in\Pi} LR_{bT}(\pi)$ when $\Pi = [0,1]$. For the location model with iid $N(0,1)$ errors, D. M. Hawkins (1977) has already investigated this behavior (heuristically). In the general model scenario considered here, this behavior is determined using the results of Theorem 4.

COROLLARY 1: *Suppose the conditions of Theorem* 4(a) (*resp.* 4(b), ...,4(e)) *and the null hypothesis* H_0 *hold. Then,* $\sup_{\pi\in[0,1]} W_T(\pi) \xrightarrow{p} \infty$ (*resp.* $\sup_{\pi\in[0,1]} LM_{aT}(\pi) \xrightarrow{p} \infty$, ..., $\sup_{\pi\in[0,1]} LR_{bT}(\pi) \xrightarrow{p} \infty$).

COMMENTS: 1. The corollary shows that the restriction in Theorem 4 to sets Π whose closure is in (0,1) is not made only for technical convenience. Unless Π is bounded away from zero and one, critical values for the test statistics $\sup_{\pi\in\Pi} W_T(\pi)$, ..., $\sup_{\pi\in\Pi} LR_{bT}(\pi)$ must diverge to infinity as $T \to \infty$ to obtain a sequence of level α tests. By bounding π away from zero and one, however, a fixed critical value suffices for all T large. This suggests that the restriction of Π to a set whose closure is in (0,1) yields significant power gains if the true change point is in Π or is close to Π. Some Monte Carlo results of Talwar (1983) and James, James, and Siegmund (1987) for the location model substantiate this result. Furthermore, the Monte Carlo results of Talwar (1983) show that the test statistic $\sup_{\pi\in\Pi} W_T(\pi)$ has much closer true and nominal sizes in the location model under non–normal errors when Π is restricted than when $\Pi = [0,1]$.

2. Suppose $\hat{\pi}$ maximizes $W_T(\pi)$, $LM_{aT}(\pi)$, $LR_{aT}(\pi)$, $LM_{bT}(\pi)$, or $LR_{bT}(\pi)$ over [0,1]. For example, in the ML Example 2, if $\hat{\pi}$ maximizes $LR_{bT}(\pi)$, then $\hat{\pi}$ is the ML estimator of π for the parameter space [0,1]. By Theorem 4 and Corollary 1, $\sup_{\pi\in[\epsilon,1-\epsilon]} W_T(\pi) = O_p(1)$ $\forall\epsilon > 0$ and $\sup_{\pi\in[0,1]} W_T(\pi) \xrightarrow{p} \infty$ under the null hypothesis and analogous results hold for $LM_{aT}(\pi)$, ..., $LR_{bT}(\pi)$. In consequence, $\hat{\pi} \xrightarrow{p} \{0,1\}$ under the null hypothesis. By symmetry, presumably, $\hat{\pi} \xrightarrow{d} \mathrm{Bern}(1/2)$,

where Bern(1/2) denotes a Bernoulli distribution with parameter 1/2. In contrast, if Π has closure in (0,1) and $Q_p(\cdot)$ has a unique maximum on Π with probability one, then $\hat{\pi} \xrightarrow{d} \operatorname{argmax}\{Q_p(\pi) : \pi \in \Pi\}$ by the continuous mapping theorem. The latter distribution has support equal to Π.

4.2. Asymptotic Local Power of the Test Statistics

Next, we consider the behavior of $\hat{\theta}(\cdot)$, $W_T(\cdot)$, etc. under sequences of local alternatives. We introduce the following assumption:

ASSUMPTION 2–ℓp: *The triangular array* $\{W_{Tt} : t \leq T, T \geq 1\}$ *is such that Assumption 2 holds but with part* (d) *replaced by*

(d) $\sup_{\pi \in \Pi} \left\| \sqrt{T}\frac{\partial}{\partial m} d(E\bar{m}_T(\theta_0, \pi, \underset{\sim}{\tau}_0), \quad \hat{\gamma}_0(\pi)) - \mu(\pi) \right\| = o_p(1)$ *for some non–random,* R^v*–valued function* μ *on* Π.

We write $\mu(\pi) = (\mu_1(\pi)', \mu_2(\pi)', \mu_3(\pi)')'$ for $\mu_r(\pi) \in R^{v_r}$, $r = 1, 2, 3$.

In many cases, $\mu(\pi)$ can be expressed more simply. For example, suppose (i) $d(\bar{m}_T(\theta,\pi,\underset{\sim}{\tau}), \gamma)$ is of the form $\bar{m}_T(\theta,\pi,\underset{\sim}{\tau})'\gamma\bar{m}_T(\theta,\pi,\underset{\sim}{\tau})/2$, (ii) $\gamma_0(\pi) = \operatorname{diag}\{D_1(\pi), D_2(\pi), D_3(\pi)\}$ where $D_1(\pi)$ and $D_2(\pi)$ are $v_1 \times v_1$ matrices and $D_3(\pi)$ is a $v_3 \times v_3$ matrix, (iii) Assumption 2–ℓp holds, (iv) $\{W_{Tt} : t \leq T, T \geq 1\}$ is such that $Em_{1Tt}(W_{Tt}, \theta_{10} + \eta(t/T)/\sqrt{T}, \theta_{30}, \tau_0) = \underset{\sim}{0}$ $\forall t \leq T$, $\forall T \geq 1$ for some bounded R^p–valued function $\eta(\cdot)$ on [0,1] that is Riemann integrable on $[0,\pi]$ uniformly over $\pi \in \Pi \cup \{1\}$,[8] and (v) $\max_{t \leq T} \left\| E\frac{\partial}{\partial \theta_1'} m_{1Tt}(W_{Tt}, \theta_{10}, \theta_{30}, \tau_0) - M \right\| \to 0$ as $T \to \infty$. In this case, one can show that

$$(4.6)\qquad \mu(\pi) = \begin{bmatrix} \mu_1(\pi) \\ \mu_2(\pi) \\ \mu_3(\pi) \end{bmatrix} = \begin{bmatrix} -D_1(\pi)M\int_0^{\pi}\eta(s)\,ds \\ -D_2(\pi)M\int_{\pi}^{1}\eta(s)\,ds \\ D_3(\pi)\lim_{T\to\infty}\frac{1}{T}\Sigma_1^T Em_{3t}(\theta_{10},\theta_{30},\tau_0) \end{bmatrix}.$$

EXAMPLE 1 (GMM, cont.): In this example, we have

$$(4.7)\qquad \mu(\pi) = \begin{bmatrix} a_0'a_0\ \frac{1}{\pi}\ \lim_{T\to\infty}\frac{1}{\sqrt{T}}\Sigma_1^{T\pi}\ \ Ef(W_{Tt},\theta_{10}) \\ a_0'a_0\frac{1}{1-\pi}\lim_{T\to\infty}\frac{1}{\sqrt{T}}\Sigma_{T\pi+1}^{T}Ef(W_{Tt},\theta_{10}) \end{bmatrix}.$$

Now, suppose $\{W_{Tt} : t \le T, T \ge 1\}$ is such that $Ef(W_{Tt}, \theta_{10} + \eta(t/T)/\sqrt{T}) = \underset{\sim}{0}$ $\forall t \le T$, $\forall T \ge 1$ for $\eta(\cdot)$ as above and $\max_{t\le T}\left\|E\frac{\partial}{\partial\theta_1'}f(W_{Tt},\theta_{10}) - M\right\| \to 0$ as $T\to\infty$, where $M = \lim_{T\to\infty}\frac{1}{T}\Sigma_1^T E\frac{\partial}{\partial\theta_1'}f(W_{Tt},\theta_{10})$. Then, as in (4.6), we have

$$(4.8)\qquad \mu(\pi) = \begin{bmatrix} \mu_1(\pi) \\ \mu_2(\pi) \end{bmatrix} = \begin{bmatrix} -a_0'a_0\ \frac{1}{\pi}\ M\int_0^{\pi}\eta(s)ds \\ -a_0'a_0\frac{1}{1-\pi}M\int_{\pi}^{1}\eta(s)ds \end{bmatrix}.$$

EXAMPLE 2 (ML, cont.): In the ML case, we have

$$(4.9)\qquad \mu(\pi) = \begin{bmatrix} \lim_{T\to\infty} -\frac{1}{\sqrt{T}}\Sigma_1^{T\pi}\ \ E\frac{\partial}{\partial\theta_1}\log f_t(\theta_{10},\theta_{30}) \\ \lim_{T\to\infty} -\frac{1}{\sqrt{T}}\Sigma_{T\pi+1}^{T}E\frac{\partial}{\partial\theta_1}\log f_t(\theta_{10},\theta_{30}) \\ \lim_{T\to\infty} -\frac{1}{\sqrt{T}}\Sigma_1^{T}\ \ E\frac{\partial}{\partial\theta_3}\log f_t(\theta_{10},\theta_{30}) \end{bmatrix}.$$

If $\{W_{Tt} : t \le T, T \ge 1\}$ is such that $E\frac{\partial}{\partial\theta_1}\log f_t(\theta_{10} + \eta(t/T)/\sqrt{T}, \theta_{30}) = \underset{\sim}{0}$ $\forall t \le T$, $\forall T \ge 1$ for $\eta(\cdot)$ as above and $\max_{t\le T}\left\|-E\frac{\partial^2}{\partial\theta_1\partial\theta_1'}\log f_t(\theta_{10},\theta_{30}) - M\right\| \to 0$, where $M = \lim_{T\to\infty} -\frac{1}{T}\Sigma_1^T E\frac{\partial^2}{\partial\theta_1\partial\theta_1'}\log f_t(\theta_{10},\theta_{30})$, then

$$(4.10)\qquad \mu(\pi) = \begin{bmatrix} -M\int_0^{\pi}\eta(s)\,ds \\ -M\int_{\pi}^{1}\eta(s)\,ds \\ \lim_{T\to\infty} -\frac{1}{\sqrt{T}}\,\Sigma_1^T E\frac{\partial}{\partial\theta_3}\log f_t(\theta_{10},\theta_{30}) \end{bmatrix}.$$

THEOREM 5: *Suppose Assumptions* 2–ℓp, 3, *and* 4 *hold. Given any set* Π *whose closure lies in* (0,1), *the following processes indexed by* $\pi\in\Pi$ *satisfy:*

(a) $\sqrt{T}(\hat\theta(\cdot)-\theta_0) \Rightarrow -(M(\cdot)'D(\cdot)M(\cdot))^{-1}M(\cdot)'(D(\cdot)G(\cdot)+\mu(\cdot))$,

(b) $\sup_{\pi\in\Pi}\|\hat V_r(\pi)-V_r(\pi)\| \xrightarrow{p} 0$ *for* $r = 1, 2$,

(c) $W_T(\cdot) \Rightarrow Q_p^*(\cdot) = J_p^*(\cdot)'J_p^*(\cdot)$ *and* $\sup_{\pi\in\Pi} W_T(\pi) \xrightarrow{d} \sup_{\pi\in\Pi} Q_p^*(\pi)$, *where* $J_p^*(\cdot)$

$$= \frac{B^*(\cdot) - \iota(\cdot)B^*(1)}{\iota(\cdot)^{1/2}(1-\iota(\cdot))^{1/2}} - \left[\frac{1-\iota(\cdot)}{\iota(\cdot)}\right]^{1/2} AS^{-1/2}D_1^{-1}(\cdot)\mu_1(\cdot)$$
$$+ \left[\frac{\iota(\cdot)}{1-\iota(\cdot)}\right]^{1/2} AS^{-1/2}D_2^{-1}(\cdot)\mu_2(\cdot),$$

(d) $LM_{aT}(\cdot) \Rightarrow Q_p^*(\cdot)$ *and* $\sup_{\pi\in\Pi} LM_{aT}(\pi) \xrightarrow{d} \sup_{\pi\in\Pi} Q_p^*(\pi)$ *provided Assumption* 5a *also holds,*

(e) $LR_{aT}(\cdot) \Rightarrow Q_p^*(\cdot)$ *and* $\sup_{\pi\in\Pi} LR_{aT}(\pi) \xrightarrow{d} \sup_{\pi\in\Pi} Q_p^*(\pi)$ *provided Assumptions* 5a *and* 6a *hold in place of* 3,

(f) $LM_{bT}(\cdot) \Rightarrow Q_p^*(\cdot)$ *and* $\sup_{\pi\in\Pi} LM_{bT}(\pi) \xrightarrow{d} \sup_{\pi\in\Pi} Q_p^*(\pi)$ *provided Assumptions* 5b *and* 6b(i) *also hold, and*

(g) $LR_{bT}(\cdot) \Rightarrow Q_p^*(\cdot)$ *and* $\sup_{\pi\in\Pi} LR_{bT}(\pi) \xrightarrow{d} \sup_{\pi\in\Pi} Q_p^*(\pi)$ *provided Assumptions* 5b *and* 6b *hold in place of* 3,

where $G(\cdot)$ *is as in Theorem* 2, $B^*(\cdot)$ *and* $\iota(\cdot)$ *are as in Theorem* 4, $\mu(\cdot) = (\mu_1(\cdot)', \mu_2(\cdot)', \mu_3(\cdot)')'$ *is as in Assumption* 2–ℓp, $A = I_p$ *when* $p = v_1$, $A = (C(\pi)C(\pi)')^{-1/2}C(\pi)$ *for any* $\pi\in\Pi$ *when* $p < v_1$ *for* $C(\pi) = (M'D_1(\pi)M)^{-1}MD_1(\pi)S^{1/2}$ (*since* $(C(\pi)C(\pi)')^{-1/2}C(\pi)$ *does not depend on* π

in this case by Assumption 4(iv)), *and* $B^*(\cdot) = AB_1(\cdot)$. *Furthermore, the convergence in* (a)–(g) *holds jointly.*

COMMENTS: 1. When $\mu(\cdot)$ satisfies (4.6), the limit process $Q_p^*(\cdot)$ of Theorem 5 depends on $\eta(\cdot)$ in the following way: $Q_p^*(\cdot) = J_p^*(\cdot)' J_p^*(\cdot)$ and

$$J_p^*(\cdot) = \frac{B^*(\cdot) - \iota(\cdot)B^*(1)}{\iota(\cdot)^{1/2}(1-\iota(\cdot))^{1/2}}$$

$$+ AS^{-1/2}M\left[\left[\frac{1-\iota(\cdot)}{\iota(\cdot)}\right]^{1/2}\int_0^{\iota(\cdot)} \eta(s)ds - \left[\frac{\iota(\cdot)}{1-\iota(\cdot)}\right]^{1/2}\int_{\iota(\cdot)}^1 \eta(s)ds\right] . \tag{4.11}$$

2. For fixed $\pi \in (0,1)$, $Q_p^*(\pi)$ has a noncentral chi–square distribution with p degrees of freedom and noncentrality parameter given by the squared length of the sum of the last two summands that define $J_p^*(\pi)$ in Theorem 5(c).

EXAMPLE 1 (GMM, cont.): When $\mu(\cdot)$ is as in (4.8), the limit process $Q_p^*(\cdot) = J_p^*(\cdot)' J_p^*(\cdot)$ of Theorem 5 is given by (4.11).

EXAMPLE 2 (ML, cont.): In the ML case, the limit process $Q_p^*(\cdot) = J_p^*(\cdot)' J_p^*(\cdot)$ of Theorem 5 simplifies when (4.10) holds as follows:

$$J_p^*(\cdot) = \frac{B^*(\cdot) - \iota(\cdot)B^*(1)}{\iota(\cdot)^{1/2}(1-\iota(\cdot))^{1/2}}$$

$$- M^{1/2}\left[\left[\frac{1-\iota(\cdot)}{\iota(\cdot)}\right]^{1/2}\int_0^{\iota(\cdot)} \eta(s)ds - \left[\frac{\iota(\cdot)}{1-\iota(\cdot)}\right]^{1/2}\int_{\iota(\cdot)}^1 \eta(s)ds\right] , \tag{4.12}$$

since $A = I_p$, $S = M$, and $D_r(\pi) = I_p$ for $r = 1, 2$.

The local power results of Theorem 5 can be used to show that the tests based on $\sup_{\pi\in\Pi} W_T(\pi), \ldots, \sup_{\pi\in\Pi} LR_{bT}(\pi)$ each have non–trivial power against alternatives for which the θ_{1t} parameter is non–constant on Π . These results are analogous to results obtained by Ploberger *et al.* (1989, Cor. 1) for the fluctuation test in the more restrictive context of testing for pure structural change in an iid linear regression model.

COROLLARY 2: *Suppose the assumptions of Theorem* 5(c) (*resp.* 5(d)–5(g)) *hold with* $\mu(\cdot)$ *as in* (4.6) *but with* $\eta(\cdot)$ *replaced by* $\xi\eta(\cdot)$. *If* η *is not almost everywhere* (*Lebesgue*) *equal to a constant vector on* Π , *then*

$$\lim_{\xi\to\infty}\lim_{T\to\infty} P(\sup_{\pi\in\Pi} W_T(\pi) > c_\alpha) = 1$$

(*resp.* $\lim_{\xi\to\infty}\lim_{T\to\infty} P(\sup_{\pi\in\Pi} LM_{aT}(\pi) > c_\alpha) = 1, \ldots, \lim_{\xi\to\infty}\lim_{T\to\infty} P(\sup_{\pi\in\Pi} LR_{bT}(\pi) > c_\alpha) = 1$), *where* c_α *is as defined above and* $\alpha \in (0,1)$.

Next, using Theorem 5, we can establish a weak optimality result for the test statistics $\sup_{\pi\in\Pi} W_T(\pi), \ldots, \sup_{\pi\in\Pi} LR_{bT}(\pi)$ for testing against the alternatives in $\cup_{\pi\in\Pi} H_1(\pi)$. This result is a generalization to multiparameter two-sided tests of a result of Davies (1977, Thm. 4.2) for scalar parameter one–sided tests. The result shows that as the significance level α goes to zero, the power against all local alternatives of the level α test based on $\sup_{\pi\in\Pi} W_T(\pi)$ is at least as large as that of the level α test based on $W_T(\tilde\pi)$ for any fixed $\tilde\pi \in \Pi$. Thus, if $W_T(\tilde\pi)$ possesses asymptotic local power optimality properties against certain alternatives, e.g., as it does in the ML case against one–time structural changes (i.e., for $\eta(s) = 0$ for $s < \tilde\pi$, $= \delta$ for $s \geq \tilde\pi$), then $\sup_{\pi\in\Pi} W_T(\pi)$ inherits these same properties as $\alpha \to 0$. The same results also hold for $\sup_{\pi\in\Pi} LM_{aT}(\pi), \ldots, \sup_{\pi\in\Pi} LR_{bT}(\pi)$.

THEOREM 6: *Let* η *denote a bounded* R^p*–valued function on* [0,1] *that is Riemann integrable on* $[0,\pi]$ *uniformly over* $\pi \in \Pi \cup \{1\}$. *Let* Ξ *denote the set of all such functions* η *for which there exists a distribution* P_η *of the triangular array* $\{W_{Tt} : t \leq T, T \geq 1\}$ *such that Assumptions* 2–ℓp, 3, *and* 4 *hold with* $\mu(\cdot)$ *as in* (4.6). *Then,*

$$\lim_{\alpha\to 0}\ \inf_{\eta\in\Xi}\ \inf_{\tilde\pi\in\Pi}\ \lim_{T\to\infty}\Big[P_\eta(\sup_{\pi\in\Pi} W_T(\pi) > c_\alpha) - P_\eta(W_T(\tilde\pi) > \tilde c_\alpha)\Big] \geq 0\ , \tag{4.13}$$

where the critical values c_α *and* $\tilde c_\alpha$ *are such that the tests based on* $\sup_{\pi\in\Pi} W_T(\pi)$ *and*

$W_T(\tilde{\pi})$ *have asymptotic level* $\alpha \in (0,1)$. *The result* (4.13) *also holds with* $W_T(\cdot)$ *replaced by* (i) $LM_{aT}(\cdot)$, (ii) $LR_{aT}(\cdot)$, (iii) $LM_{bT}(\cdot)$, *or* (iv) $LR_{bT}(\cdot)$ *provided* Ξ *is restricted to include only functions* η *for which* P_η *also satisfies* (i) *Assumption* 5a, (ii) 5a *and* 6a *in place of* 3, (iii) 5b *and* 6b(i), *or* (iv) 5b *and* 6b *in place of* 3, *respectively*.

COMMENT: The optimality result (4.13) is referred to above as a *weak* result because it appears that α must be quite small before the result is illustrative of finite sample behavior of the test statistics $\sup_{\pi \in \Pi} W_T(\pi)$ and $W_T(\tilde{\pi})$. Nevertheless, the result does serve to indicate that as α decreases the difference decreases between the power function of the level α test based on $\sup_{\pi \in \Pi} W_T(\pi)$ and the envelope of the power functions of the level α tests based on $W_T(\tilde{\pi})$ for fixed $\tilde{\pi} \in \Pi$.

5. MONTE CARLO RESULTS

This section presents some Monte Carlo results regarding the finite sample size and power properties of the tests discussed above. It considers tests of parameter instability in a linear regression model with iid normal errors. A "sup F test," which is covered by the results of this paper, is compared with the CUSUM test of Brown, Durbin, and Evans (1975), the fluctuation test of Sen (1980) and Ploberger, Krämer, and Kontrus (1989), and the "midpoint F test" (i.e., Chow test) of one–time structural change occurring at the midpoint of the sample. The CUSUM test is the best known parameter instability test in the literature for linear regression models. The fluctuation test is a more recently developed alternative to the CUSUM test that has been shown to have some power advantages, see Krämer and Sonnberger (1986) and Kontrus (1984). The midpoint F test also has been suggested for use as a general test of parameter instability. In addition, its power properties for a one-time shift at the midpoint of the sample can be used to measure the cost in terms of power of having to estimate the change point using the sup F test, as opposed to knowing the change point.

The sup F test is the supremum over $\pi \in [.15, .85]$ of the F test statistic, $F_T(\pi)$, for testing one–time structural change in the regression parameters at point π. ($F_T(\pi)$ is also known as the Chow test statistic for testing change at time $t = T\pi$.) Since $F_T(\pi) = ((T-2p)/T)W_T(\pi)/p$ (where p is the number of regressors), the asymptotic results of Section 4 apply to the sup F test and the critical values of Table 1 divided by p are appropriate. See Krämer and Sonnberger (1986, pp. 49–53, 59–61), for example, for the definition of the CUSUM and fluctuation tests. The midpoint F test is $F_T(.5)$.

For reasons of comparability, the model used in the Monte Carlo experiment reported here is the same as that used by Krämer and Sonnberger (1986, pp. 64–69) and Kontrus (1984). It is given by

$$(5.1) \qquad Y_t = X_t'\beta_t + U_t, \; X_t = (1, (-1)^t)', \; U_t \sim \text{iid } N(0, \sigma^2), \; t = 1, \ldots, T.$$

Under the null hypothesis, $\beta_t = \beta_0 \; \forall t$ for some $\beta_0 \in R^2$ and the distributions of the four test statistics are invariant with respect to β_0 and σ^2. For simplicity, we take $\beta_0 = \underset{\sim}{0}$ and $\sigma^2 = 1$ for the significance level results.

The sup F, CUSUM, and fluctuation tests are asymptotic tests. In consequence, the first set of Monte Carlo results compares each of these tests' true and nominal significance levels, see Table 2. The midpoint F test is exact, and hence, no results are reported for it in Table 2. Nominal (asymptotic) significance levels of 1%, 5%, and 10% and sample sizes of T = 30, 60, 120, 240, and 1000 are considered. Fifty thousand repetitions are used for each of the first three sample sizes, while ten thousand repetitions are used for the latter two.

Table 2 shows that the sup F test over–rejects when T = 30. The over–rejection is fairly small for the 5% and 10% nominal significance levels, but is marked for the 1% result. For each of the larger sample sizes, however, the sup F test exhibits close agreement between its true and nominal significance levels. In all such cases except (T = 60, 1%), it under–rejects slightly, and hence, is slightly conservative. The CUSUM and fluctuation tests both exhibit marked under–rejection for sample sizes $T = 30, 60$, and 120 and all significance levels. The under–rejection of the fluctuation test in particular is quite severe for $T = 30$ and 60. In sum, for sample sizes $T = 60$ and 120, the sup F test exhibits the best performance of the three tests under the null, while for $T = 30$ no single test is dominant.

Next, we consider the power properties of the sup F, CUSUM, fluctuation, and midpoint F tests. Again for reasons of comparability, we consider more or less the same alternatives as are considered by Krämer and Sonnberger (1986) and Kontrus (1984). In particular, we consider the simple case of a single shift $\Delta\beta$ in the regression parameters:

$$\beta_t = (0,0)' \text{ for } t \leq T\pi^* \text{ and } \beta_t = \Delta\beta \text{ for } t > T\pi^* . \tag{5.2}$$

We vary (1) the time π^* of change, (2) the magnitude $\|\Delta\beta\| = b/\sqrt{T}$ of change, (3) the angle ψ between $\Delta\beta$ and $\lim_{T\to\infty} \frac{1}{T} \Sigma_1^T EX_t = (1,0)'$, and (4) the sample size T. We report results for $\pi^* = .15, .3, .5, .7, .85$; $b = 4.8, 7.2, 9.6, 12.0$; $\psi = 0^\circ, 36^\circ, 54^\circ, 90^\circ$; $\sigma^2 = 1$; and $T = 30, 60, 120$. One thousand repetitions are used for each scenario. Note that $\|\Delta\beta\| = b/\sqrt{T}$ is a decreasing function of T, and so, the power of a given test for any fixed value of b is not necessarily increasing in T.

The sup F and midpoint F tests possess several desirable invariance properties that the other two tests do not possess. First, their power is invariant with respect to the "direction of time," and hence, they have equal finite sample power for $\pi^* = .15$ and .85, as well as for $\pi^* = .3$ and .7. In contrast, the CUSUM test will be seen to have greater power against changes early in the sample as opposed to late and vice versa for the fluctuation test. Second, the sup F and midpoint F tests have power that is invariant with respect to the angle ψ. In contrast, as shown by Krämer, Ploberger, and Alt (1988), the CUSUM test has local power that depends greatly on the angle ψ. Its local power is zero for $\psi = 90^\circ$ and is maximized for $\psi = 0^\circ$. The fluctuation test also has power that depends on ψ. For this test, local power is least for $\psi = 45^\circ$ and greatest for $\psi = 0^\circ$ or 90°.

Tables 3A, 3B, and 3C report the powers of the four tests under consideration for significance level 5% and sample sizes T = 30, 60, and 120 respectively. Asymptotic critical values are used for the sup F, CUSUM, and fluctuation tests, while exact critical values are used for the midpoint F test. For T = 60 and 120, all of the tests are conservative (i.e., they do not over–reject), and hence, comparisons between the tests are meaningful. For T = 30, the sup F test over–rejects and its power performance is (misleadingly) enhanced accordingly.

Tables 4A, 4B, and 4C give the size corrected powers of the four tests under consideration for significance level 5% and sample sizes T = 30, 60, and 120 respectively. The exact critical values used for the sup F, CUSUM, and fluctuation tests were generated by Monte Carlo using 50,000 repetitions.[9] These tables eliminate the power distortions that arise due to under– or over–rejection under the null when asymptotic critical values are used. The power of the midpoint F test that is given in Tables 4A–4C, is the same as that given in Tables 3A–3C, since no size correction is needed for that test.

First, we discuss the general characteristics of the power of each of the four tests as exhibited in Tables 3A–3C and 4A–4C. Thereafter, we compare the power of the four tests. The sup F test has power that is invariant with respect to the direction of time and angle ψ, as discussed above. Its power is always greatest when change occurs in the middle of the sample $(\pi^* = .5)$ and lowest when it occurs very early or very late in the sample $(\pi^* = .15 \text{ or } .85)$.

The CUSUM test has power that declines quickly with the magnitude of the angle ψ. It has no power at angle 90° and little or no power at angle 54° in all of the scenarios considered. The CUSUM test has greater power when change occurs in the first half of the sample than when it occurs in the second half. In each scenario, its power is greatest for

$\pi^* = .3$, next greatest for $\pi^* = .15$, and next for $\pi^* = .5$. For $\pi^* = .7$ and .85, there is a large drop off in power.

The fluctuation test has maximum (and approximately equal) power at angles 0° and 90° and minimum (and approximately equal) power at 36° and 54°. The difference in power between the former and the latter is usually in the range of 10% to 50%. Thus, the fluctuation test's power varies significantly with the angle ψ , but it varies much less than does the power of the CUSUM test. The fluctuation test has greater power when change occurs in the latter half of the sample than when it occurs in the earlier half. The difference between halves, however, is not nearly as pronounced as for the CUSUM test. For the fluctuation test, the values of the change point π^* in order of declining power are .5, .7, .3, .85, .15.

The midpoint F test has power that is invariant with respect to the direction of time and the angle ψ. As with the sup F test, it has greatest power for $\pi^* = .5$ and least power for $\pi^* = .15$ and .85. The dropoff in power from $\pi^* = .5$ to $\pi^* = .3$ and .7 to $\pi^* = .15$ and .85 is sharp. The midpoint F test has very little power when $\pi^* = .15$ and .85.

We now discuss the relative power of the tests based on the asymptotic critical values (Tables 3A–3C). The symbols + and = in the tables indicate those scenarios where the CUSUM, fluctuation, and midpoint F tests have power greater than, respectively equal to, that of the sup F test. The scarcity of such symbols indicates the nearly uniform strict dominance of the sup F test over the CUSUM and fluctuation tests in these tables. The comparison between the sup F and midpoint F tests is as expected. The sup F test has higher power at each change point except the midpoint. The sup F test has much higher power than the midpoint F test for $\pi^* = .15$ and .85, but noticeably lower power at $\pi^* = .5$. In terms of overall power performance, the sup F test is clearly the best of the four tests in Tables 3A–3C.

Next, consider the size adjusted power results of Tables 4A–4C. The greatest change between Tables 3A–3C and 4A–4C is the increased power of the CUSUM and fluctuation tests due to the correction of their under–rejection. For the CUSUM test, the increase in power is not sufficient to alter the near uniform strict dominance of the sup F test. For the fluctuation test, the increase in power does alter the dominance of the sup F test. The differences between these two tests vary mainly with π^* rather than with ψ or T . For $\pi^* = .15$, the sup F test is generally significantly more powerful than the fluctuation test across different values of ψ and T . For $\pi^* = .3$, it varies from slightly more powerful to somewhat more powerful. For $\pi^* = .5$, it is somewhat less powerful. For $\pi^* = .7$, it varies from equally powerful to somewhat more powerful. For $\pi^* = .85$, it is somewhat more powerful. In sum, the overall power properties of the sup F test are preferable to those of the fluctuation test even after size correction, but they are not uniformly dominant. The comparison between the sup F test and the midpoint F test in Tables 4A–4C is quite similar to that in Tables 3A–3C.

To conclude, the sup F test is clearly the best test of the four considered in the Monte Carlo experiment in the limited scenarios considered here. It has much less nominal/true size discrepancy than the CUSUM and fluctuation tests. It has much better power than the CUSUM test with or without size correction. It has much better power than the fluctuation test without size correction and somewhat better power with size correction. It has much better power than the midpoint F test for changes away from the midpoint, but less power for changes at the midpoint.

Although the range of alternative scenarios considered here is quite limited, two factors suggest a preference for the sup F test over the CUSUM and fluctuation tests for a wider range of alternatives. First, the CUSUM test is essentially a one degree of freedom test and the resulting problems that it exhibits in the tables here for many angles will manifest themselves in other scenarios as well. Second, the considerable under–rejection of the fluctuation test will shrink its (non–size corrected) power not just for the one–time shift alternatives considered here, but for all alternatives of parameter instability.

Lastly, we report some Monte Carlo comparisons between the sup F–test and the average F–test, defined by $\int_0^1 F_T(\pi)d\pi \ \left[= \frac{1}{T}\Sigma_{t=2}^{T-2} F_T(t/T)\right]$. In terms of the distribution of weight given to $F_T(\pi)$ for different values of π, the latter is the extreme opposite of the sup F–test. In terms of closeness of true and nominal asymptotic significance levels, the sup F–test and average F–test perform quite similarly with the sup F–test being better overall by a slight margin. For all three sample sizes, the size corrected (and non–size corrected) power of these two tests is remarkably similar for significance levels 5% and 10%. The sup F–test has greater power than the average F–test when $d = .15$ or $.85$ and $b = 7.2, 9.6$, or 12.0, whereas the average F–test has greater power than the sup F–test when $d = .5$ and $b = 4.8$. For all other scenarios, the power of the two tests is very similar. On the other hand, when the significance level is 1%, the sup F–test has a slight, but clear, advantage over the average F–test in terms of size corrected (and non–size corrected) power. This result is in accord with the large sample small significance level optimality result for the sup F–test given in Theorem 6 above.

These Monte Carlo results help to alleviate the possible criticism that the justifications given above for the use of the sup function to define the test statistic of interest are somewhat weak. The reason is that the results indicate that the performance of a test that is based on a function g of $\{F_T(\pi) : \pi \in \Pi\}$ is not overly sensitive to the choice of g, at least for functions g within some class of "reasonable" functions.

APPENDIX

The proofs of Theorems 1 and 2 use the following two lemmas:

LEMMA A–1: *Let* $\{Z_t\{\xi\} : t \geq 1\}$ *be a sequence of rv's indexed by* $\xi \in \Xi$, *where* Ξ *is an arbitrary space. Suppose* $\sup_{\xi\in\Xi}\left|\frac{1}{T}\Sigma_1^T Z_t(\xi)\right| \to 0$ *a.s. Then,* $\sup_{\pi\in[\epsilon,1]}\sup_{\xi\in\Xi}\left|\frac{1}{T\pi}\Sigma_1^{T\pi} Z_t(\xi)\right| \xrightarrow{p} 0$ *and* $\sup_{\pi\in[\epsilon,1-\epsilon]}\sup_{\xi\in\Xi}\left|\frac{1}{T-T\pi}\Sigma_{T\pi+1}^{T} Z_t(\xi)\right| \xrightarrow{p} 0 \ \forall 0 < \epsilon < 1$.

PROOF OF LEMMA A–1: To prove the first result, note that

$$\text{(A.1)} \qquad \sup_{\pi\in[\epsilon,1]}\sup_{\xi\in\Xi}\left|\frac{1}{T\pi}\Sigma_1^{T\pi} Z_t(\xi)\right| \leq \sup_{S\geq T\epsilon}\sup_{\xi\in\Xi}\left|\frac{1}{S}\Sigma_1^{S} Z_t(\xi)\right| .$$

By a characterization of a.s. convergence, the right–hand side (rhs) of (A.1) $\xrightarrow{p} 0$ iff $\sup_{\xi\in\Xi}\left|\frac{1}{T\epsilon}\Sigma_1^{T\epsilon} Z_t(\xi)\right| \to 0$ a.s. (e.g., see Chow and Teicher (1978, Lemma 3.3.1, p. 66)). The latter holds by assumption. The second result of the lemma is proved in a similar fashion by writing $\Sigma_{T\pi+1}^{T} = \Sigma_1^T - \Sigma_1^{T\pi}$. □

LEMMA A–2: *Suppose* $\hat\theta(\pi)$ *minimizes a random real function* $Q_T(\theta,\pi)$ *over* $\theta \in \Theta \subset R^{2p+p_3}$ *for each* $\pi \in \Pi \subset [0,1]$ *with probability* $\to 1$. *If*

(a) $\sup_{\pi\in\Pi,\,\theta\in\Theta}|Q_T(\theta,\pi) - Q(\theta,\pi)| \xrightarrow{p} 0$ *for some real function* Q *on* $\Theta \times \Pi$ *and*

(b) *for every neighborhood* Θ_0 $(\subset \Theta)$ *of* θ_0, $\inf_{\pi\in\Pi}(\inf_{\theta\in\Theta/\Theta_0} Q(\theta,\pi) - Q(\theta_0,\pi)) > 0$,

then $\sup_{\pi\in\Pi}\|\hat\theta(\pi) - \theta_0\| \xrightarrow{p} 0$.

PROOF OF LEMMA A–2: By Assumption (b), given any neighborhood Θ_0 of θ_0, there exists a constant $\delta > 0$ such that $\inf_{\tilde\pi\in\Pi}\left[\inf_{\theta\in\Theta/\Theta_0} Q(\theta,\tilde\pi) - Q(\theta_0,\tilde\pi)\right] \geq \delta > 0$. Thus,

$$P\left[\hat\theta(\pi) \in \Theta/\Theta_0 \text{ for some } \pi \in \Pi\right] \leq P\left[\inf_{\tilde\pi\in\Pi}(Q(\hat\theta(\pi),\tilde\pi) - Q(\theta_0,\tilde\pi)) \geq \delta \text{ for some } \pi \in \Pi\right]$$

$$\text{(A.2)} \qquad \leq P\left[Q(\hat\theta(\pi),\pi) - Q(\theta_0,\pi) \geq \delta \text{ for some } \pi \in \Pi\right] \to 0 ,$$

where " $\to 0$ " holds provided $\sup_{\pi\in\Pi}|Q(\hat\theta(\pi),\pi) - Q(\theta_0,\pi)| \xrightarrow{p} 0$. Using Assumptions (a) and (b), the latter follows from

$$
\begin{aligned}
(\text{A.3}) \qquad 0 &\le \inf_{\pi\in\Pi}\Big[Q(\hat\theta(\pi),\pi) - Q(\theta_0,\pi)\Big] \le \sup_{\pi\in\Pi}\Big[Q(\hat\theta(\pi),\pi) - Q(\theta_0,\pi)\Big] \\
&\le \sup_{\pi\in\Pi}\Big[Q(\hat\theta(\pi),\pi) - Q_T(\hat\theta(\pi),\pi)\Big] + \sup_{\pi\in\Pi}\Big[Q_T(\hat\theta(\pi),\pi) - Q(\theta_0,\pi)\Big] \\
&\le \sup_{\pi\in\Pi}\Big[Q(\hat\theta(\pi),\pi) - Q_T(\hat\theta(\pi),\pi)\Big] + \sup_{\pi\in\Pi}\Big[Q_T(\theta_0,\pi) - Q(\theta_0,\pi)\Big] \\
&\le 2 \sup_{\pi\in\Pi,\,\theta\in\Theta} |Q_T(\theta,\pi) - Q(\theta,\pi)| \xrightarrow{p} 0\,. \;\square
\end{aligned}
$$

PROOF OF THEOREM 1: We show that Assumption 1 implies that conditions (a) and (b) of Lemma A–2 hold with $Q_T(\theta,\pi) = d(\bar{m}_T(\theta,\pi,\hat{\underset{\sim}{\tau}}(\pi)), \hat\gamma(\pi))$ and $Q(\theta,\pi) = d(m(\theta,\pi,\underset{\sim}{\tau}_0), \gamma_0(\pi))$. Condition (b) of Lemma A–2 holds by Assumption 1(d).

To obtain condition (a), note that Lemma A–1 and Assumption 1(b) imply that

$$
(\text{A.4}) \qquad \sup_{\pi\in\Pi,\theta\in\Theta,\underset{\sim}{\tau}\in\mathcal{T}_0\times\mathcal{T}_0} \|\bar{m}_T(\theta,\pi,\underset{\sim}{\tau}) - m(\theta,\pi,\underset{\sim}{\tau})\| \xrightarrow{p} 0\,.
$$

Condition (a) of Lemma A–2 now follows from

$$
\begin{aligned}
(\text{A.5}) \qquad &\sup_{\pi\in\Pi,\,\theta\in\Theta} |d(\bar{m}_T(\theta,\pi,\hat{\underset{\sim}{\tau}}(\pi)),\hat\gamma(\pi)) - d(m(\theta,\pi,\underset{\sim}{\tau}_0),\gamma_0(\pi))| \\
\le\; &\sup_{\pi\in\Pi,\,\theta\in\Theta} |d(\bar{m}_T(\theta,\pi,\hat{\underset{\sim}{\tau}}(\pi)),\hat\gamma(\pi)) - d(m(\theta,\pi,\hat{\underset{\sim}{\tau}}(\pi)),\hat\gamma(\pi))| \\
+\; &\sup_{\pi\in\Pi,\,\theta\in\Theta} |d(m(\theta,\pi,\hat{\underset{\sim}{\tau}}(\pi)),\hat\gamma(\pi)) - d(m(\theta,\pi,\underset{\sim}{\tau}_0),\gamma_0(\pi))| \\
\xrightarrow{p}\; &0\,,
\end{aligned}
$$

where " $\xrightarrow{p} 0$ " uses Assumptions 1(a) and (c) and (A.4). $\square$

For notational simplicity, we use the following abbreviations below: $\bar{m}_T(\theta,\pi)$ for $\bar{m}_T(\theta,\pi,\hat{\underset{\sim}{\tau}}(\pi))$, $m(\theta,\pi)$ for $m(\theta,\pi,\underset{\sim}{\tau}_0)$, $\hat\gamma$ for $\hat\gamma(\pi)$, γ_0 for $\gamma_0(\pi)$, $\hat{\underset{\sim}{\tau}}$ for $\hat{\underset{\sim}{\tau}}(\pi)$, $\tilde\theta_a$ for $\tilde\theta_a(\pi)$, and $\tilde\theta_b$ for $\tilde\theta_b(\pi)$.

PROOF OF THEOREM 2: Element by element mean value expansions of $\sqrt{T}\frac{\partial}{\partial\theta}d(\bar{m}_T(\hat{\theta}(\pi), \pi), \hat{\gamma})$ about θ_0 give: $\forall j = 1, \ldots, 2p + p_3$,

$$
\begin{aligned}
o_{p\pi}(1) &= \sqrt{T}\frac{\partial}{\partial\theta_j}d(\bar{m}_T(\hat{\theta}(\pi), \pi), \hat{\gamma}) \\
&= \sqrt{T}\frac{\partial}{\partial\theta_j}d(\bar{m}_T(\theta_0, \pi), \hat{\gamma}) + \frac{\partial^2}{\partial\theta'\partial\theta_j}d(\bar{m}_T(\theta^*,\pi), \hat{\gamma})\sqrt{T}(\hat{\theta}(\pi) - \theta_0) ,
\end{aligned} \tag{A.6}
$$

where $\theta = (\theta_1, \ldots, \theta_{2p+p_3})'$ and θ^* $(= \theta^*(\pi))$ is on the line segment joining $\hat{\theta}(\pi)$ and θ_0. The latter property and Assumption 2(a) imply that $\theta^* = \theta_0 + o_{p\pi}(1)$. The first equality of (A.6) holds because $\hat{\theta}(\pi)$ minimizes $d(\bar{m}_T(\theta,\pi), \hat{\gamma})$ and $\hat{\theta}(\pi)$ is in the interior of Θ $\forall\pi \in \Pi$ with probability $\to 1$ by Assumption 2(a).

Below we show that

$$
\frac{\partial^2}{\partial\theta'\partial\theta_j}d(\bar{m}_T(\theta^*,\pi), \hat{\gamma}) = \frac{\partial^2}{\partial\theta'\partial\theta_j}d(m(\theta_0, \pi, \underset{\sim}{\tau}_0), \gamma_0) + o_{p\pi}(1) , \tag{A.7}
$$

where $m(\theta,\pi,\underset{\sim}{\tau}) = \left[\pi m_1(\theta_1, \theta_3, \tau_1)', (1-\pi)m_1(\theta_2, \theta_3, \tau_2)', \pi m_3(\theta_1, \theta_3, \tau_1)' + (1-\pi)m_3(\theta_2, \theta_3, \tau_2)'\right]'$ and $\frac{\partial^2}{\partial\theta\partial\theta'}d(m(\theta_0, \pi, \underset{\sim}{\tau}_0), \gamma_0) = M(\pi)'D(\pi)M(\pi)$. Also, we show that

$$
\sqrt{T}\frac{\partial}{\partial\theta}d(\bar{m}_T(\theta_0, \cdot), \hat{\gamma}(\cdot)) \Rightarrow M(\cdot)'D(\cdot)G(\cdot) \tag{A.8}
$$

as a process indexed by $\pi \in \Pi$. These results, equation (A.6), Assumption 2(h), and the continuous mapping theorem combine to give the desired result:

$$
\begin{aligned}
\sqrt{T}(\hat{\theta}(\cdot) - \theta_0) &= -(M(\cdot)'D(\cdot)M(\cdot))^{-1}\sqrt{T}\frac{\partial}{\partial\theta}d(\bar{m}_T(\theta_0, \cdot), \hat{\gamma}(\cdot)) + o_{p\pi}(1) \\
&\Rightarrow -(M(\cdot)'D(\cdot)M(\cdot))^{-1}M(\cdot)'D(\cdot)G(\cdot) .
\end{aligned} \tag{A.9}
$$

To show (A.7), we proceed as follows: $\forall j, \ell = 1, \ldots, 2p + p_3$,

$$
\begin{aligned}
\frac{\partial^2}{\partial\theta_j\partial\theta_\ell}d(\bar{m}_T(\theta^*,\pi), \hat{\gamma}) &= \frac{\partial^2}{\partial\theta_j\partial\theta_\ell}\bar{m}_T(\theta^*,\pi)'\frac{\partial}{\partial m}d(\bar{m}_T(\theta^*,\pi), \hat{\gamma}) \\
&\quad + \frac{\partial}{\partial\theta_j}\bar{m}_T(\theta^*,\pi)'\frac{\partial^2}{\partial m\partial m'}d(\bar{m}_T(\theta^*,\pi), \hat{\gamma})\frac{\partial}{\partial\theta_\ell}\bar{m}_T(\theta^*,\pi) .
\end{aligned} \tag{A.10}
$$

By Lemma A–1 and Assumption 2(g),

$$\text{(A.11)} \qquad \sup_{\pi\in\Pi,\theta_1\in\Theta_{10},\theta_3\in\Theta_{30},\tau_1\in\mathcal{T}_0} \left\| \frac{1}{T}\Sigma_1^{T\pi}(m_{rt}(\theta_1,\theta_3,\tau_1) - Em_{rt}(\theta_1,\theta_3,\tau_1)) \right\| \xrightarrow{p} 0$$

for $r = 1, 3$. Similarly, this result with $\Sigma_1^{T\pi}$ replaced by $\Sigma_{T\pi+1}^{T}$ also holds. Combining these results gives

$$\text{(A.12)} \qquad \sup_{\pi\in\Pi,\theta\in\Theta_0,\underset{\sim}{\tau}\in\mathcal{T}_0\times\mathcal{T}_0} \|\bar{m}_T(\theta,\pi,\underset{\sim}{\tau}) - E\bar{m}_T(\theta,\pi,\underset{\sim}{\tau})\| \xrightarrow{p} 0 .$$

Equation (A.12) and Assumptions 2(a), (b), (c), and (g) give

$$\text{(A.13)} \qquad \begin{aligned} \sup_{\pi\in\Pi}\|\bar{m}_T(\theta^*,\pi,\hat{\underset{\sim}{\tau}}) - m(\theta_0,\pi,\underset{\sim}{\tau}_0)\| &\le \sup_{\pi\in\Pi}\|\bar{m}_T(\theta^*,\pi,\hat{\underset{\sim}{\tau}}) - E\bar{m}_T(\theta^*,\pi,\hat{\underset{\sim}{\tau}})\| \\ &+ \sup_{\pi\in\Pi}\|E\bar{m}_T(\theta^*,\pi,\hat{\underset{\sim}{\tau}}) - m(\theta^*,\pi,\hat{\underset{\sim}{\tau}})\| + \sup_{\pi\in\Pi}\|m(\theta^*,\pi,\hat{\underset{\sim}{\tau}}) - m(\theta_0,\pi,\underset{\sim}{\tau}_0)\| \xrightarrow{p} 0 , \end{aligned}$$

where $E\bar{m}_T(\theta^*,\pi,\hat{\underset{\sim}{\tau}}) = E\bar{m}_T(\theta,\pi,\underset{\sim}{\tau})\Big|_{(\theta,\underset{\sim}{\tau})=(\theta^*,\hat{\underset{\sim}{\tau}})}$. The second summand on the rhs $\xrightarrow{p} 0$, because the limit $\lim_{T\to\infty}\frac{1}{T\pi}\Sigma_1^{T\pi}Em_{rt}(\theta_1,\theta_3,\tau_1)$ exists not only uniformly over $\Theta_{10}\times\Theta_{30}\times\mathcal{T}_0$, but also uniformly over $\Theta_{10}\times\Theta_{30}\times\mathcal{T}_0\times\Pi$ for $r = 1, 3$. Using (A.13), the uniform continuity of $\frac{\partial}{\partial m}d(m,\gamma)$, and Assumption 2(b), we get

$$\text{(A.14)} \qquad \frac{\partial}{\partial m}d(\bar{m}_T(\theta^*,\pi),\hat{\gamma}) = \frac{\partial}{\partial m}d(m(\theta_0,\pi,\underset{\sim}{\tau}_0),\gamma_0) + o_{p\pi}(1) = \underset{\sim}{0} + o_{p\pi}(1) ,$$

where the second equality holds by Assumptions 2(b), (c), (d), (f), and (g).

Using Assumptions 2(c) and (g) and Markov's inequality, one obtains $\frac{\partial^2}{\partial\theta_j\partial\theta_\ell}\bar{m}_T(\theta^*,\pi) = O_{p\pi}(1)$ $\forall j, \ell = 1, \ldots, 2p + p_3$. This result and (A.14) imply that the first term on the rhs of (A.10) is $o_{p\pi}(1)$.

Uniform continuity of $\frac{\partial^2}{\partial m\partial m'}d(m,\gamma)$, (A.13), and Assumption 2(b) give

$$\text{(A.15)} \qquad \frac{\partial^2}{\partial m\partial m'}d(\bar{m}_T(\theta^*,\pi),\hat{\gamma}) = \frac{\partial^2}{\partial m\partial m'}d(m(\theta_0,\pi,\underset{\sim}{\tau}_0),\gamma_0) + o_{p\pi}(1) = D(\pi) + o_{p\pi}(1) .$$

It follows from Lemma A–1 and Assumption 2(g) that

$$\text{(A.16)}\qquad \begin{gathered} \sup_{\pi\in\Pi,\theta\in\Theta_0,\underset{\sim}{\tau}\in\mathcal{T}_0\times\mathcal{T}_0}\left\|\frac{\partial}{\partial\theta'}\bar{m}_T(\theta,\pi,\underset{\sim}{\tau})-E\frac{\partial}{\partial\theta'}\bar{m}_T(\theta,\pi,\underset{\sim}{\tau})\right\| \xrightarrow{p} 0 \text{ and} \\ \sup_{\pi\in\Pi,\theta\in\Theta_0,\underset{\sim}{\tau}\in\mathcal{T}_0\times\mathcal{T}_0}\left\|E\frac{\partial}{\partial\theta'}\bar{m}_T(\theta,\pi,\underset{\sim}{\tau})-\lim_{T\to\infty}E\frac{\partial}{\partial\theta'}\bar{m}_T(\theta,\pi,\underset{\sim}{\tau})\right\| \to 0\,. \end{gathered}$$

These results, Assumptions 2(a), (c), and (g) and the definition $M(\pi)=\lim_{T\to\infty}E\frac{\partial}{\partial\theta'}\bar{m}_T(\theta_0,\pi,\underset{\sim}{\tau}_0)$ give

$$\text{(A.17)}\qquad \frac{\partial}{\partial\theta'}\bar{m}_T(\theta^*,\pi)=M(\pi)+o_{p\pi}(1)\,.$$

Equations (A.15) and (A.17) imply that the second term of (A.10) equals $[M(\pi)'D(\pi)M(\pi)]_{j\ell}+o_{p\pi}(1)$, and hence, (A.7) is established.

To establish (A.8), we write

$$\text{(A.18)}\qquad \sqrt{T}\frac{\partial}{\partial\theta}d(\bar{m}_T(\theta_0,\pi),\hat{\gamma})=\sqrt{T}\left[\frac{\partial}{\partial\theta'}\bar{m}_T(\theta_0,\pi)\right]'\frac{\partial}{\partial m}d(\bar{m}_T(\theta_0,\pi),\hat{\gamma})\,.$$

As in (A.17), we have

$$\text{(A.19)}\qquad \frac{\partial}{\partial\theta'}\bar{m}_T(\theta_0,\pi)=M(\pi)+o_{p\pi}(1)\,.$$

By the mean value theorem, the j–th element of $\sqrt{T}\frac{\partial}{\partial m}d(\bar{m}_T(\theta_0,\pi),\hat{\gamma})$ $j=1,\ldots,v$ can be expanded about $E\bar{m}_T(\theta_0,\pi,\underset{\sim}{\tau}_0)$ to get:

$$\text{(A.20)}\qquad \begin{aligned} &\sqrt{T}\frac{\partial}{\partial m_j}d(\bar{m}_T(\theta_0,\pi,\hat{\underset{\sim}{\tau}}),\hat{\gamma}) \\ &=\sqrt{T}\frac{\partial}{\partial m_j}d(E\bar{m}_T(\theta_0,\pi,\underset{\sim}{\tau}_0),\hat{\gamma})+\frac{\partial^2}{\partial m'\partial m_j}d(m^*,\hat{\gamma})\sqrt{T}(\bar{m}_T(\theta_0,\pi,\hat{\underset{\sim}{\tau}})-E\bar{m}_T(\theta_0,\pi,\underset{\sim}{\tau}_0))\,, \end{aligned}$$

where m^* depends on π and is on the line segment joining $\bar{m}_T(\theta_0,\pi,\hat{\underset{\sim}{\tau}})$ and $E\bar{m}_T(\theta_0,\pi,\underset{\sim}{\tau}_0)$. By (A.13), $m^*=m(\theta_0,\pi,\underset{\sim}{\tau}_0)+o_{p\pi}(1)$.

The first term of the rhs of (A.20) is $o_{p\pi}(1)$ by Assumption 2(d). Also, using (A.13) and Assumptions 2(b) and (f),

$$\text{(A.21)}\qquad \frac{\partial^2}{\partial m'\partial m_j}d(m^*,\hat{\gamma})=[D(\pi)]_j+o_{p\pi}(1)\,,$$

where $[D(\pi)]_j$ denotes the j–th row of $D(\pi)$.

Using (A.18)–(A.21), the proof of (A.8) is complete once we show that

$$\text{(A.22)} \qquad \sqrt{T}(\bar{m}_T(\theta_0, \pi, \hat{\tau}) - E\bar{m}_T(\theta_0, \pi, \tau_0)) \Rightarrow G(\cdot)\,.$$

A mean value expansion of the j–th element of $\bar{m}_T(\theta_0, \pi, \hat{\tau}(\pi))$ about τ_0 yields

$$\begin{aligned} &\sqrt{T}(\bar{m}_{Tj}(\theta_0, \pi, \hat{\tau}(\pi)) - E\bar{m}_{Tj}(\theta_0, \pi, \tau_0)) \\ \text{(A.23)} \qquad &= \sqrt{T}(\bar{m}_{Tj}(\theta_0, \pi, \tau_0) - E\bar{m}_{Tj}(\theta_0, \pi, \tau_0)) + \frac{\partial}{\partial \tau'}\bar{m}_{Tj}(\theta_0, \pi, \tau^*)\sqrt{T}(\hat{\tau}(\pi) - \tau_0) \\ &= \sqrt{T}(\bar{m}_{Tj}(\theta_0, \pi, \tau_0) - E\bar{m}_{Tj}(\theta_0, \pi, \tau_0)) + o_{p\pi}(1)\,, \end{aligned}$$

where τ^* depends on π and lies on the line segment joining $\hat{\tau}(\pi)$ and τ_0. The second equality of (A.23) holds using Assumptions 2(c) and (g), since $\frac{\partial}{\partial \tau'}\bar{m}_T(\theta_0, \pi, \tau^*)$ $= \left[\pi dm_1(\theta_{10}, \theta_{30}, \tau_0)', (1-\pi)dm_1(\theta_{10}, \theta_{30}, \tau_0)', dm_3(\theta_{10}, \theta_{30}, \tau_0)'\right]' + o_{p\pi}(1) \quad = 0$ $+ o_{p\pi}(1)$. Stacking (A.23) for $j = 1, \ldots, 2p + p_3$ and using Assumption 2(e) gives (A.22). □

PROOF OF THEOREM 3: First, suppose $\hat{V}_r(\pi)$ is as defined in (2.15) and (2.16). By the argument of (A.15), $\sup_{\pi\in\Pi}\|\hat{D}_r(\pi) - D_r(\pi)\| = o_p(1)$ for $r = 1, 2$ and by the argument of (A.16) and (A.17), $\sup_{\pi\in\Pi}\|\hat{M}_r(\pi) - M\| = o_p(1)$ for $r = 1, 2$. (Note that (A.16) still holds with the terms in $\|\cdot\|$ multiplied by $1/\pi$ or $1/(1-\pi)$.) These results and Assumptions 2(h) and 3 give the desired result. The proof for the case where $V_r(\pi)$ is defined as in (2.15) and (2.17) is analogous. □

The following Lemma is used in Sections 3.1 and 3.2 to obtain simple sufficient conditions for Assumptions 4(ii) and 5a(ii) to hold.

LEMMA A–3: *Let* $M(\pi)$ *be a square nonsingular matrix of the form*
$\begin{bmatrix} \pi M & 0 & \pi M_{13} \\ 0 & (1-\pi)M & (1-\pi)M_{13} \\ \pi M_{31} & (1-\pi)M_{31} & M_{33} \end{bmatrix}$, *where* $\pi \in (0,1)$, $M \in R^{p\times p}$, *and* $M_{33} \in R^{p_3\times p_3}$.

Let $G = (G_1', G_2', G_3')'$ *be any vector in* R^{2p+p_3} *and let* $H = [I_p \vdots -I_p \vdots \underset{\sim}{0}]$. *Let* $M_*(\pi) = \begin{bmatrix} \pi M & \underset{\sim}{0} \\ \underset{\sim}{0} & (1-\pi)M \end{bmatrix}$, $G_* = (G_1', G_2')'$, *and* $H_* = [I_p \vdots -I_p]$. *Then,* $HM(\pi)^{-1}G = H_*M_*(\pi)^{-1}G_*$.

PROOF OF LEMMA A–3: Let $v = (v_1', v_2', v_3')' = M(\pi)^{-1}G$ and $\tilde{v} = (\tilde{v}_1', \tilde{v}_2')' = M_*(\pi)^{-1}G_*$. Since $M(\pi)v = G$, we have

$$\begin{bmatrix} \pi M & \underset{\sim}{0} \\ \underset{\sim}{0} & (1-\pi)M \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} - \begin{bmatrix} \pi M_{13} v_3 \\ (1-\pi)M_{13}v_3 \end{bmatrix},$$

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} \tilde{v}_1 \\ \tilde{v}_2 \end{bmatrix} - \begin{bmatrix} M^{-1}M_{13}v_3 \\ M^{-1}M_{13}v_3 \end{bmatrix}, \text{ and} \tag{A.24}$$

$$HM(\pi)^{-1}G = v_1 - v_2 = \tilde{v}_1 - \tilde{v}_2 = H_*M_*(\pi)^{-1}G_* . \quad \square$$

PROOF OF THEOREM 4: First we establish part (a). By Assumption 4(iii),

$$C(\pi) = (M'D_1(\pi)M)^{-1}M'D_1(\pi)S^{1/2} = (M'D_2(\pi)M)^{-1}M'D_2(\pi)S^{1/2} . \tag{A.25}$$

Throughout this proof, we use the subscript $*$ as a deletion operator as in Assumption 4 (except for $\mathcal{J}_*(\pi)$ which is defined below). By Theorem 2, Assumption 4(ii), and (A.25),

$$\begin{aligned} \sqrt{T}(\hat{\theta}_1(\cdot) - \hat{\theta}_2(\cdot)) &= H\sqrt{T}(\hat{\theta}(\cdot) - \theta_0) \\ &\Rightarrow H(M(\cdot)'D(\cdot)M(\cdot))^{-1}M(\cdot)'D(\cdot)G(\cdot) \\ &= H_*(M_*(\cdot)'D_*(\cdot)M_*(\cdot))^{-1}M_*(\cdot)'D_*(\cdot)G_*(\cdot) \\ &= C(\cdot)\left[\tfrac{1}{\iota(\cdot)}B_1(\cdot) - \tfrac{1}{1-\iota(\cdot)}(B_1(1) - B_1(\cdot))\right], \end{aligned} \tag{A.26}$$

where $G_*(\cdot) = (\nu_1(\cdot)', \nu_1(1)' - \nu_1(\cdot)')'$ and $\nu_1(\cdot) = S^{1/2}B_1(\cdot)$.

By Theorem 3 and (A.25),

$$\hat{V}_1(\cdot) + \hat{V}_2(\cdot) \Rightarrow V_1(\cdot) + V_2(\cdot) = \tfrac{1}{\iota(\cdot)}C(\cdot)C(\cdot)' + \tfrac{1}{1-\iota(\cdot)}C(\cdot)C(\cdot)' . \tag{A.27}$$

Equations (A.26) and (A.27) and the continuous mapping theorem give

$$\text{(A.28)}\quad \begin{aligned} W_T(\cdot) \Rightarrow\ & (B_1(\cdot) - \iota(\cdot)B_1(1))'C(\cdot)'(C(\cdot)C(\cdot)')^{-1}C(\cdot) \\ & \times(B_1(\cdot) - \iota(\cdot)B_1(1))/[\iota(\cdot)(1-\iota(\cdot))] . \end{aligned}$$

Now, if $p = v_1$, then $C(\pi)'(C(\pi)C(\pi)')^{-1}C(\pi)$ equals I_p $\forall \pi \in \Pi$, since the former is a projection matrix onto the full p–dimensional space. In this case, the first result of part (a) holds by (A.28) with $B_1(\cdot) = B^*(\cdot)$.

If $p < v_1$, then let $B^*(\cdot) = (C(\cdot)C(\cdot)')^{-1/2}C(\cdot)B_1(\cdot)$. Since $A = (C(\pi)C(\pi)')^{-1/2}C(\pi)$ does not depend on π by Assumption 4(iv), we have $B^*(\cdot) = AB_1(\cdot)$ and $AA' = I_p$. Thus, $B^*(\cdot)$ is a p–vector of independent Brownian motions and the first result of part (a) holds by (A.28) with $B^*(\cdot) = AB_1(\cdot)$. The second and third results of part (a) follow from the first using the continuous mapping theorem. The same is true in parts (b)–(e), so it suffices to establish the first result in parts (b)–(e).

Next we establish part (b). Standard arguments give: For $r = 1, 2$,

$$\text{(A.29)}\quad \tilde{M}_r(\pi) = M + o_{p\pi}(1)\,,\ \tilde{D}_r(\pi) = D_r(\pi) + o_{p\pi}(1)\,,\ \text{and}\ \tilde{V}_r(\pi) = V_r(\pi) + o_{p\pi}(1)\,.$$

In addition, we show below that

$$\text{(A.30)}\quad \sqrt{T}\frac{\partial}{\partial\theta}d(\bar{m}_T(\tilde{\theta}_a, \pi), \hat{\gamma}) = O_{p\pi}(1)\,.$$

Hence, it suffices to show that $LM^0_{aT}(\cdot) \Rightarrow Q_p(\cdot)$, where $LM^0_{aT}(\pi)$ is the same as $LM_{aT}(\pi)$, defined in (3.8), but with M, $D_r(\pi)$, and $V_r(\pi)$ in place of $\tilde{M}_r(\pi)$, $\tilde{D}_r(\pi)$, and $\tilde{V}_r(\pi)$, respectively.

Let $\Omega(\pi) = \text{Var}(G(\pi))$ and $\mathcal{J}(\pi) = M(\pi)'D(\pi)M(\pi)$. Then, $\Omega_*(\pi) = \begin{bmatrix} \pi S & \underset{\sim}{0} \\ \underset{\sim}{0} & (1-\pi)S \end{bmatrix}$. Let $\mathcal{J}_*(\pi)$ denote $M_*(\pi)'D_*(\pi)M_*(\pi)$. Let $\doteq$ denote equality that holds with probability $\to 1$. We have

$$
\begin{aligned}
\mathrm{LM}^0_{\mathrm{aT}}(\pi) &= \mathrm{T}\frac{\partial}{\partial(\theta_1', \theta_2')}\mathrm{d}(\bar{\mathrm{m}}_{\mathrm{T}}(\tilde{\theta}_{\mathrm{a}},\pi),\ \hat{\gamma})\mathcal{J}_*^{-1}\mathrm{H}_*'\left[\mathrm{H}_*\mathcal{J}_*^{-1}\mathrm{M}_*(\pi)'\mathrm{D}_*\Omega_*\mathrm{D}_*\mathrm{M}_*(\pi)\mathcal{J}_*^{-1}\mathrm{H}_*'\right]^{-1} \\
&\quad \times \mathrm{H}_*\mathcal{J}_*^{-1}\frac{\partial}{\partial(\theta_1', \theta_2')'}\mathrm{d}(\bar{\mathrm{m}}_{\mathrm{T}}(\tilde{\theta}_{\mathrm{a}},\pi),\hat{\gamma}) \\
&\doteq \mathrm{T}\tilde{\lambda}'\mathrm{H}_*\mathcal{J}_*^{-1}\mathrm{H}_*'\left[\mathrm{H}\mathcal{J}^{-1}\mathrm{M}(\pi)'\mathrm{D}\Omega\mathrm{D}\mathrm{M}(\pi)\mathcal{J}^{-1}\mathrm{H}'\right]^{-1}\mathrm{H}_*\mathcal{J}_*^{-1}\mathrm{H}_*'\tilde{\lambda} \\
&= \mathrm{T}\tilde{\lambda}'\mathrm{H}\mathcal{J}^{-1}\mathrm{H}'\left[\mathrm{H}\mathcal{J}^{-1}\mathrm{M}(\pi)'\mathrm{D}\Omega\mathrm{D}\mathrm{M}(\pi)\mathcal{J}^{-1}\mathrm{H}'\right]^{-1}\mathrm{H}\mathcal{J}^{-1}\mathrm{H}'\tilde{\lambda} \qquad \text{(A.31)} \\
&\doteq \mathrm{T}\frac{\partial}{\partial\theta'}\mathrm{d}(\bar{\mathrm{m}}_{\mathrm{T}}(\tilde{\theta}_{\mathrm{a}},\pi),\ \hat{\gamma})\mathcal{J}^{-1}\mathrm{H}'\left[\mathrm{H}\mathcal{J}^{-1}\mathrm{M}(\pi)\mathrm{D}\Omega\mathrm{D}\mathrm{M}(\pi)\mathcal{J}^{-1}\mathrm{H}'\right]^{-1} \\
&\quad \times \mathrm{H}\mathcal{J}^{-1}\frac{\partial}{\partial\theta}\mathrm{d}(\bar{\mathrm{m}}_{\mathrm{T}}(\tilde{\theta}_{\mathrm{a}},\pi),\ \hat{\gamma})\ ,
\end{aligned}
$$

where the second and fourth equalities use the fact that $\frac{\partial}{\partial\theta}\mathrm{d}(\bar{\mathrm{m}}_{\mathrm{T}}(\tilde{\theta}_{\mathrm{a}},\pi),\ \hat{\gamma}) = -\mathrm{H}'\tilde{\lambda}$ for some p–vector of Lagrange multipliers $\tilde{\lambda}$ $(= \tilde{\lambda}(\pi))$, the second equality also uses Assumption 4(ii), the third equality uses Assumption 5a(ii), and the dependence on π of $\tilde{\theta}_{\mathrm{a}}$, $\hat{\gamma}$, $\mathcal{J}_*$, D_*, Ω_*, $\tilde{\lambda}$, $\mathcal{J}$, D, and Ω is suppressed for notational simplicity.

Now we determine the asymptotic distribution of the rhs of (A.31) viewed as a process indexed by $\pi \in \Pi$. Element by element mean value expansions about θ_0 yield: $\forall \mathrm{j} = 1,\ \ldots,\ 2\mathrm{p} + \mathrm{p}_3$,

$$
\text{(A.32)} \quad \sqrt{\mathrm{T}}\frac{\partial}{\partial\theta_{\mathrm{j}}}\mathrm{d}(\bar{\mathrm{m}}_{\mathrm{T}}(\tilde{\theta}_{\mathrm{a}},\pi),\hat{\gamma}) = \sqrt{\mathrm{T}}\frac{\partial}{\partial\theta_{\mathrm{j}}}\mathrm{d}(\bar{\mathrm{m}}_{\mathrm{T}}(\theta_0,\pi),\hat{\gamma}) + \frac{\partial^2}{\partial\theta'\partial\theta_{\mathrm{j}}}\mathrm{d}(\bar{\mathrm{m}}_{\mathrm{T}}(\theta^+,\pi),\hat{\gamma})\sqrt{\mathrm{T}}(\tilde{\theta}_{\mathrm{a}} - \theta_0)\ ,
$$

where θ^+ $(= \theta^+(\pi))$ lies on the line segment joining $\tilde{\theta}_{\mathrm{a}}(\pi)$ and θ_0, and hence, satisfies $\theta^+ = \theta_0 + o_{\mathrm{p}\pi}(1)$. We stack (A.32) for $\mathrm{j} = 1,\ \ldots,\ 2\mathrm{p} + \mathrm{p}_3$ and write it as

$$
\text{(A.33)} \qquad \sqrt{\mathrm{T}}\frac{\partial}{\partial\theta}\mathrm{d}(\bar{\mathrm{m}}_{\mathrm{T}}(\tilde{\theta}_{\mathrm{a}},\pi),\ \hat{\gamma}) = \sqrt{\mathrm{T}}\frac{\partial}{\partial\theta}\mathrm{d}(\bar{\mathrm{m}}_{\mathrm{T}}(\theta_0,\pi),\ \hat{\gamma}) + \mathcal{J}^0\sqrt{\mathrm{T}}(\tilde{\theta}_{\mathrm{a}} - \theta_0)\ ,
$$

where by standard arguments $\mathcal{J}^0$ $(= \mathcal{J}^0(\pi))$ satisfies $\mathcal{J}^0(\pi) = \mathcal{J}(\pi) + o_{\mathrm{p}\pi}(1)$.

By definition of $\tilde{\theta}_{\mathrm{a}}$, we have

$$
\text{(A.34)} \qquad \underset{\sim}{0} = \mathrm{H}\sqrt{\mathrm{T}}(\tilde{\theta}_{\mathrm{a}} - \theta_0)\ .
$$

By (A.8), $\sqrt{\mathrm{T}}\frac{\partial}{\partial\theta}\mathrm{d}(\bar{\mathrm{m}}_{\mathrm{T}}(\theta_0,\cdot),\ \hat{\gamma}(\cdot)) \Rightarrow \mathrm{M}(\cdot)'\mathrm{D}(\cdot)\mathrm{G}(\cdot)$. Also, using the nonsingularity of $\mathcal{J}(\pi)$ (see Assumption 2(h)), we get $\mathcal{J}^0(\pi)^{-1}\mathcal{J}^0(\pi) \doteq \mathrm{I}_{2\mathrm{p}+\mathrm{p}_3}$. Pre–multiplication of (A.33) by $H\mathcal{J}^0(\cdot)^{-1}$ now gives

$$\begin{aligned}
&H\mathcal{J}^0(\cdot)^{-1}\sqrt{T}\tfrac{\partial}{\partial\theta}d(\bar{m}_T(\tilde{\theta}_a,\cdot),\ \hat{\gamma}(\cdot)) \\
&= H\mathcal{J}^0(\cdot)^{-1}\sqrt{T}\tfrac{\partial}{\partial\theta}d(\bar{m}_T(\theta_0,\cdot),\ \hat{\gamma}(\cdot)) + H\mathcal{J}^0(\cdot)^{-1}\mathcal{J}^0(\cdot)\sqrt{T}(\tilde{\theta}_a(\cdot)-\theta_0) \\
\text{(A.35)}\quad &\doteq H\mathcal{J}^0(\cdot)^{-1}\sqrt{T}\tfrac{\partial}{\partial\theta}d(\bar{m}_T(\theta_0,\cdot),\ \hat{\gamma}(\cdot)) \\
&\Rightarrow H\mathcal{J}(\cdot)^{-1}M(\cdot)'D(\cdot)G(\cdot) \\
&= H_*\mathcal{J}_*(\cdot)^{-1}M_*(\ )'D_*(\cdot)G_*(\cdot)\,,
\end{aligned}$$

using Assumption 4(ii). Equations (A.30) and (A.35) yield

$$\begin{aligned}
&H\mathcal{J}(\cdot)^{-1}\sqrt{T}\tfrac{\partial}{\partial\theta}d(\bar{m}_T(\tilde{\theta}_a,\cdot),\ \hat{\gamma}(\cdot)) \Rightarrow H_*\mathcal{J}_*(\cdot)^{-1}M_*(\cdot)'D_*(\cdot)G_*(\cdot) \\
\text{(A.36)}\quad &\qquad = \frac{1}{\iota(\cdot)(1-\iota(\cdot))}C(\cdot)(B_1(\cdot)-\iota(\cdot)B_1(1))\,.
\end{aligned}$$

In addition, Assumption 4(ii) implies that

$$\begin{aligned}
&H\mathcal{J}^{-1}M(\cdot)'D\Omega DM(\cdot)\mathcal{J}^{-1}H' = H_*\mathcal{J}_*^{-1}M_*(\cdot)'D_*\Omega_*D_*M_*(\cdot)\mathcal{J}_*^{-1}H_*' \\
\text{(A.37)}\quad &\qquad = V_1(\cdot)+V_2(\cdot) = \frac{1}{\iota(\cdot)(1-\iota(\cdot))}C(\cdot)C(\cdot)'\,.
\end{aligned}$$

By (A.31), $LM^0_{aT}(\pi)$ is a quadratic form in the vector given in (A.36) with weight matrix given by the inverse of the matrix in (A.37). Hence, using (A.36) and (A.37), $LM^0_{aT}(\cdot)$ has the same limit as that of $W_T(\cdot)$ in (A.28). As above, this limit is $Q_p(\cdot)$.

For part (b), it remains to show (A.30). With probability $\to 1$, $\tilde{\theta}_a(\pi)$ is in the interior of Θ $\forall\pi\in\Pi$ and there exists a rv $\tilde{\lambda}(\pi)$ of Lagrange multipliers such that

$$\text{(A.38)}\qquad \tfrac{\partial}{\partial\theta}d(\bar{m}_T(\tilde{\theta}_a,\pi),\ \hat{\gamma}) + H'\tilde{\lambda}(\pi) = \underset{\sim}{0}\quad \forall\pi\in\Pi\,.$$

Equations (A.35) and (A.38) combine to give

$$\text{(A.39)}\qquad -H\mathcal{J}^0(\pi)^{-1}H'\sqrt{T}\tilde{\lambda}(\pi) \doteq H\mathcal{J}^0(\pi)^{-1}\sqrt{T}\tfrac{\partial}{\partial\theta}d(\bar{m}_T(\tilde{\theta}_a,\pi),\ \hat{\gamma}) = O_{p\pi}(1)\,.$$

Since $H\mathcal{J}^0(\pi)^{-1}H' = H\mathcal{J}(\pi)^{-1}H' + o_{p\pi}(1)$ and $H\mathcal{J}(\pi)^{-1}H'$ is nonsingular, equations (A.38) and (A.39) imply that $\sqrt{T}\tilde{\lambda}(\pi) = O_{p\pi}(1)$ and that (A.30) holds.

We now prove part (c). Suppose that Assumption 6a holds. A two–term Taylor expansion of $d(\bar{m}_T(\tilde{\theta}_a,\pi), \hat{\gamma})$ about $\hat{\theta}$ $(= \hat{\theta}(\pi))$ gives

$$
\begin{aligned}
LR_{aT}(\pi) &= 2T\Big[d(\bar{m}_T(\tilde{\theta}_a,\pi), \hat{\gamma}) - d(\bar{m}_T(\hat{\theta},\pi),\hat{\gamma})\Big]/\hat{b} \\
&= 2T\frac{\partial}{\partial\theta'}d(\bar{m}_T(\hat{\theta},\pi),\hat{\gamma})(\tilde{\theta}_a-\hat{\theta})/\hat{b} + T(\tilde{\theta}_a-\hat{\theta})'\frac{\partial^2}{\partial\theta\partial\theta'}d(\bar{m}_T(\theta_\times,\pi),\hat{\gamma})(\tilde{\theta}_a-\hat{\theta})/\hat{b} \qquad \text{(A.40)} \\
&\doteq T(\tilde{\theta}_a - \hat{\theta})'\mathcal{J}_\times(\tilde{\theta}_a - \hat{\theta})/\hat{b} ,
\end{aligned}
$$

where $\theta_\times$ $(= \theta_\times(\pi))$ lies on the line segment joining $\tilde{\theta}_a$ and $\hat{\theta}$, and hence, $\theta_\times = \theta_0 + o_{p\pi}(1)$, $\mathcal{J}_\times$ $(= \mathcal{J}_\times(\pi))$ is defined implicitly and satisfies $\mathcal{J}_\times(\pi) = \mathcal{J}(\pi) + o_{p\pi}(1)$, and " $\doteq$ " holds by the first order conditions for the estimator $\hat{\theta}$.

Applying the mean value theorem element by element and stacking the equations yields

$$
\begin{aligned}
\sqrt{T}\frac{\partial}{\partial\theta}d(\bar{m}_T(\tilde{\theta}_a, \pi), \hat{\gamma}) &= \sqrt{T}\frac{\partial}{\partial\theta}d(\bar{m}_T(\hat{\theta},\pi), \hat{\gamma}) + \mathcal{J}_\dagger\sqrt{T}(\tilde{\theta}_a - \hat{\theta}) \qquad \text{(A.41)} \\
&\doteq \mathcal{J}_\dagger\sqrt{T}(\tilde{\theta}_a - \hat{\theta})
\end{aligned}
$$

for a matrix $\mathcal{J}_\dagger$ $(= \mathcal{J}_\dagger(\pi))$ that satisfies $\mathcal{J}_\dagger(\pi) = \mathcal{J}(\pi) + o_{p\pi}(1)$. Pre–multiplying (A.41) by $\mathcal{J}_\dagger^{-1}$ and substituting the result in (A.40) gives

$$
\begin{aligned}
LR_{aT}(\pi) &\doteq T\frac{\partial}{\partial\theta'}d(\bar{m}_T(\tilde{\theta}_a, \pi), \hat{\gamma})\mathcal{J}_\dagger^{-1}\mathcal{J}_\times\mathcal{J}_\dagger^{-1}\frac{\partial}{\partial\theta}d(\bar{m}_T(\tilde{\theta}_a, \pi), \hat{\gamma})/\hat{b} \\
&= T\tilde{\lambda}'H\mathcal{J}^{-1}H'\tilde{\lambda}/b + o_{p\pi}(1) \\
&= T\tilde{\lambda}'H_*\mathcal{J}_*^{-1}H_*'\Big[H_*\mathcal{J}_*^{-1}H_*'\Big]^{-1}H_*\mathcal{J}_*^{-1}H_*'\tilde{\lambda}/b + o_{p\pi}(1) \qquad \text{(A.42)} \\
&= T\frac{\partial}{\partial(\theta_1', \theta_2')}d(\bar{m}_T(\tilde{\theta}_a,\pi),\hat{\gamma})\mathcal{J}_*^{-1}H_*'\Big[H_*\mathcal{J}_*^{-1}M_*(\pi)'D_*\Omega_*D_*M_*(\pi)\mathcal{J}_*^{-1}H_*'\Big]^{-1} \\
&\quad \times H_*\mathcal{J}_*^{-1}\frac{\partial}{\partial(\theta_1', \theta_2')'}d(\bar{m}_T(\tilde{\theta}_a,\pi),\hat{\gamma}) + o_{p\pi}(1) \\
&= LM_{aT}^0(\pi) + o_{p\pi}(1) ,
\end{aligned}
$$

where the second equality uses (A.38), the third equality uses Assumption 5a(ii), the fifth equality uses (A.31), and the fourth equality uses the fact that by Assumption 6a,

$$
M_*(\pi)'D_*(\pi)\Omega_*(\pi)D_*(\pi)M_*(\pi) = bM_*(\pi)'D_*(\pi)M_*(\pi) = b\mathcal{J}_*(\pi) \;\; \forall\pi \in \Pi . \qquad \text{(A.43)}
$$

Part (c) follows from (A.42) and the proof above that $LM^0_{aT}(\cdot) \Rightarrow Q_p(\cdot)$.

The proof of part (d) is the same as that of part (b) given in equations (A.29) to (A.39) with the following changes: $D_r(\pi) \to I_p$, $V_r(\pi) \to \delta_r^{-1}M^{-1}SM^{-1}$, $\tilde{\theta}_a \to \tilde{\theta}_b$, $LM^0_{aT}(\pi) \to LM^0_{bT}(\pi)$, $LM_{aT}(\pi) \to LM_{bT}(\pi)$, $\mathcal{J}(\pi) \to M(\pi)$, $M(\pi)'D\Omega DM(\pi) \to \Omega$, $d(\bar{m}_T(\tilde{\theta}_a, \pi), \hat{\gamma}) \to \bar{\rho}_T(\tilde{\theta}_b, \pi, \hat{\underset{\sim}{\tau}}(\pi))$, $\mathcal{J}^0 \to M^0(\pi)$, where $M^0(\pi) = M(\pi) + o_{p\pi}(1)$, and $M(\cdot)'D(\cdot)G(\cdot) \to G(\cdot)$. Analogous changes are made to the quantities with a subscript $*$. Note that with these changes, the third equality of (A.31) holds by Assumption 4(ii) (which in this case says that $HM(\pi)^{-1}x = H_*M_*(\pi)^{-1}x_*$ $\forall x \in R^{2p+p_3}$, $\forall \pi \in \Pi$), and hence, no analogue of Assumption 5a(ii) is needed in the proof of part (d). The assertion following (A.34) that $\sqrt{T}\frac{\partial}{\partial\theta}\bar{\rho}_T(\theta_0, \cdot, \hat{\underset{\sim}{\tau}}(\cdot)) \Rightarrow G(\cdot)$ is verified by noting that

$$
\begin{aligned}
\sqrt{T}\frac{\partial}{\partial\theta}\bar{\rho}_T(\theta_0, \cdot, \hat{\underset{\sim}{\tau}}(\cdot)) &= \sqrt{T}(\bar{m}_T(\theta_0, \cdot, \hat{\underset{\sim}{\tau}}(\cdot)) - E\bar{m}_T(\theta_0, \cdot, \underset{\sim}{\tau}_0)) \\
&\quad + \sqrt{T}\frac{\partial}{\partial m}d(E\bar{m}_T(\theta_0, \cdot, \underset{\sim}{\tau}_0), \gamma_0(\cdot)) \\
&\Rightarrow G(\cdot),
\end{aligned}
\tag{A.44}
$$

by Assumptions 2(d) and (e) and (A.23).

The proof of part (e) is the same as that of part (c) given in (A.40) to (A.43) with the same changes as in the previous paragraph plus the following changes: $LR_{aT}(\pi) \to LR_{bT}(\pi)$, $\hat{b} \to \hat{c}$, $\mathcal{J}_\times \to M_\times(\pi)$, where $M_\times(\pi) = M(\pi) + o_{p\pi}(1)$, $\mathcal{J}_\dagger \to M_\dagger(\pi)$, where $M_\dagger(\pi) = M(\pi) + o_{p\pi}(1)$, $\mathcal{J} \to M(\pi)$, $b \to c$, and Assumption 6a $\to$ Assumption 6b. □

PROOF OF COROLLARY 1: The process $BB(\cdot) = B^*(\cdot) - \iota(\cdot)B^*(1)$, which appears in the definition of $Q_p(\cdot)$, is a p–vector of independent Brownian bridge processes on [0,1]. An alternative method of defining such a process is via a p–vector $BM(\cdot)$ of independent Brownian motion processes on $[0,\infty)$. In particular, we have

$$
\{BB(\pi) : \pi \in [0,1]\} \equiv \{(1-\pi)BM(\pi/(1-\pi)) : \pi \in [0,1]\},
\tag{A.45}
$$

where $\equiv$ denotes equality in distribution. Hence, we have

$$
\begin{aligned}
P\Big[\sup_{\pi\in[\pi_1,\pi_2]} Q_p(\pi) < c\Big] &= P\Big[\sup_{\pi\in[\pi_1,\pi_2]} BM\Big[\tfrac{\pi}{1-\pi}\Big]' BM\Big[\tfrac{\pi}{1-\pi}\Big] \Big/ \Big[\tfrac{\pi}{1-\pi}\Big] < c\Big] \\
&= P\Big[\sup_{s\in[1,\pi_2(1-\pi_1)/(\pi_1(1-\pi_2))]} BM\Big[\tfrac{\pi_1 s}{1-\pi_1}\Big]' BM\Big[\tfrac{\pi_1 s}{1-\pi_1}\Big] \Big/ \Big[\tfrac{\pi_1 s}{1-\pi_1}\Big] < c\Big] \qquad \text{(A.46)} \\
&= P\Big[\sup_{s\in[1,\pi_2(1-\pi_1)/(\pi_1(1-\pi_2))]} \overline{BM}(s)'\overline{BM}(s)/s < c\Big]
\end{aligned}
$$

for all $0 < \pi_1 \le \pi_2 < 1$ and $c \ge 0$, where the second equality holds by change of variables with $s = \Big[\frac{1-\pi_1}{\pi_1}\Big]\frac{\pi}{1-\pi}$, $\overline{BM}(s) = BM\Big[\frac{\pi_1 s}{1-\pi_1}\Big] \Big/ \Big[\frac{\pi_1}{1-\pi_1}\Big]^{1/2}$ by definition, and $\overline{BM}(\cdot)$ is also a Brownian motion on $[0,\infty)$ (by direct verification).

The result of Corollary 1 is now obtained as follows:

$$
\begin{aligned}
\lim_{T\to\infty} P\Big[\sup_{\pi\in[0,1]} W_T(\pi) < c\Big] &\le \overline{\lim_{\epsilon\to 0}} \lim_{T\to\infty} P\Big[\sup_{\pi\in[\epsilon,1-\epsilon]} W_T(\pi) < c\Big] \\
&= \overline{\lim_{\epsilon\to 0}} P\Big[\sup_{\pi\in[\epsilon,1-\epsilon]} Q_p(\pi) < c\Big] = \overline{\lim_{\epsilon\to 0}} P\Big[\sup_{s\in[1,(1-\epsilon)^2/\epsilon^2]} BM(s)'BM(s)/s < c\Big] \qquad \text{(A.47)} \\
&= P\Big[\sup_{s\in[1,\infty)} BM(s)'BM(s)/s\Big] = 0\,,
\end{aligned}
$$

where the first equality holds by Theorem 4, the second by (A.46), and the last by well-known properties of Brownian motion (i.e., the law of the iterated logarithm). The proof is identical for $LM_{aT}(\pi), \ldots, LR_{bT}(\pi)$. □

PROOF OF THEOREM 5: Part (a) holds by the proof of Theorem 2, noting that (A.8) holds with $M(\cdot)'D(\cdot)G(\cdot)$ replaced by $M(\cdot)'(D(\cdot)G(\cdot) + \mu(\cdot))$, since the first term of the rhs of (A.20) has probability limit $\mu(\cdot)$ rather than $\underset{\sim}{0}$ under Assumption 2–ℓp. Part (b) holds by the proof of Theorem 3.

Parts (c)–(g) hold using the proof of Theorem 4 with references to Theorems 2 and 3 replaced by references to Theorem 5(a) and (b), with $D(\cdot)G(\cdot)$ and $D_*(\cdot)G_*(\cdot)$ changed to $D(\cdot)G(\cdot) + \mu(\cdot)$ and $D_*(\cdot)G_*(\cdot) + \mu_*(\cdot)$, with the rhs of (A.26) changed to

$$C(\cdot)\Big[\tfrac{1}{\iota(\cdot)}B_1(\cdot) - \tfrac{1}{1-\iota(\cdot)}(B_1(1) - B_1(\cdot)) + \tfrac{1}{\iota(\cdot)}S^{-1/2}D_1^{-1}(\cdot)\mu_1(\cdot)$$

(A.48)

$$- \tfrac{1}{1-\iota(\cdot)}S^{-1/2}D_2^{-1}(\cdot)\mu_2(\cdot)\Big] ,$$

and with the rhs of (A.28) and (A.36) changed accordingly. □

PROOF OF COROLLARY 2: By Theorem 5(c)–(g) and the nonsingularity of $AS^{-1/2}M$ in (4.11), it suffices for Corollary 2 to show that

$$\text{(A.49)} \qquad \left[\tfrac{1-\pi}{\pi}\right]^{1/2}\int_0^\pi \eta(s)ds = \left[\tfrac{\pi}{1-\pi}\right]^{1/2}\int_\pi^1 \eta(s)ds \quad \forall\pi \in \Pi$$

does not hold. Note that (A.49) holds iff

$$\text{(A.50)} \qquad \int_0^\pi \eta_j(s)ds = \pi\int_0^1 \eta_j(s)ds \quad \forall\pi \in \Pi\ ,\ \forall j = 1, \ldots, p\ ,$$

where $\eta(\pi) = (\eta_1(\pi), \ldots, \eta_p(\pi))'$. Thus, it suffices to show that (A.50) does not hold.

Suppose (A.50) holds. Then, since $\pi\int_0^1 \eta_j(s)ds$ is twice differentiable in π $\forall\pi \in \Pi$, $\forall j = 1, \ldots, p$, so must be $\int_0^\pi \eta_j(s)ds$. In particular, we must have

$$\text{(A.51)} \qquad \frac{d}{d\pi}\int_0^\pi \eta_j(s)ds = \int_0^1 \eta_j(s)ds \ \text{ and } \ \frac{d^2}{d\pi^2}\int_0^\pi \eta_j(s)ds = 0 \quad \forall\pi \in \Pi\ ,\ \forall j = 1, \ldots, p\ .$$

This implies that $\eta_j = c_j$ almost everywhere (Lebesgue) on Π for some constant c_j $\forall j = 1, \ldots, p$, which is a contradiction. □

PROOF OF THEOREM 6: Let $u_\alpha = c_\alpha^{1/2}$ and $t_\alpha = \tilde{c}_\alpha^{1/2}$. We will show that

$$\text{(A.52)} \qquad u_\alpha - t_\alpha \to 0 \ \text{ as } \ \alpha \to 0\ .$$

Then, using Theorem 5, we have

$$\lim_{\alpha\to 0}\ \inf_{\eta\in\Xi}\ \inf_{\tilde\pi\in\Pi}\ \lim_{T\to\infty}\Big[P_\eta(\sup_{\pi\in\Pi} W_T(\pi) > c_\alpha) - P_\eta(W_T(\tilde\pi) > \tilde c_\alpha)\Big]$$

$$= \lim_{\alpha\to 0}\ \inf_{\eta\in\Xi}\ \inf_{\tilde\pi\in\Pi}\Big[P_\eta(\sup_{\pi\in\Pi} Q_p^*(\pi)^{1/2} > u_\alpha) - P_\eta(Q_p^*(\tilde\pi)^{1/2} > t_\alpha)\Big]$$

$$\text{(A.53)} \qquad \geq \lim_{\alpha\to 0}\ \inf_{\eta\in\Xi}\ \inf_{\tilde\pi\in\Pi}\Big[P_\eta(Q_p^*(\tilde\pi)^{1/2} > u_\alpha) - P_\eta(Q_p^*(\tilde\pi)^{1/2} > t_\alpha)\Big]$$

$$= 0\,,$$

where the last equality uses (A.52) and the fact that $Q_p^*(\tilde\pi)$ is a noncentral chi–square rv and the density of the square root of a noncentral chi–square rv is bounded above uniformly over all possible values of its noncentrality parameter.

To show (A.52) we use an argument similar to that of van Zwet and Oosterhoff (1967, p. 675). Let $\pi_1 = \inf\{\pi \in \Pi\} > 0$, let $\pi_2 = \sup\{\pi \in \Pi\} < 1$, and let v_α be such that $P\Big[\sup_{\pi\in[\pi_1,\pi_2]} Q_p^*(\pi)^{1/2} > v_\alpha\Big] = \alpha$. Since $t_\alpha \leq u_\alpha \leq v_\alpha$, to establish (A.52) it suffices to show that $v_\alpha - t_\alpha \to 0$ as $\alpha \to 0$.

By a result of James, James, and Siegmund (1987, eqn. (26), p. 78), we have

$$\text{(A.54)}\quad P\Big[\sup_{\pi\in[\pi_1,\pi_2]} Q_p(\pi)^{1/2} > v_\alpha\Big] = K_p v_\alpha^{p-2}\exp(-v_\alpha^2/2)\{(v_\alpha^2 - p)\log\lambda + 4 + o(1)\}$$

as $\alpha \to 0$, where $Q_p(\cdot)$ is as in Theorem 4, K_p is a constant that depends only on the dimension p of the Brownian bridge vector that underlies $Q_p(\cdot)$, and $\lambda = \pi_2(1-\pi_1)/[\pi_1(1-\pi_2)]$. Taking $\pi_1 = \pi_2 = \tilde\pi$ in (A.54) yields $\log\lambda = 0$ and

$$\text{(A.55)}\qquad P\Big[Q_p(\tilde\pi)^{1/2} > t_\alpha\Big] = K_p t_\alpha^{p-2}\exp(-t_\alpha^2/2)\{4 + o(1)\} \text{ as } \alpha \to 0\,.$$

The left–hand sides of (A.54) and (A.55) each equal α. Thus, the logs of the rhs of (A.54) and (A.55) can be equated to yield

$$(p-2)\log v_\alpha - v_\alpha^2/2 + \log\{(v_\alpha^2 - p)\log\lambda + 4 + o(1)\}$$

$$\text{(A.56)}\qquad = (p-2)\log t_\alpha - t_\alpha^2/2 + \log\{4 + o(1)\} \text{ and}$$

$$
\begin{aligned}
v_\alpha - \frac{t_\alpha^2}{v_\alpha} &= \frac{2}{v_\alpha}\Big[(p-2)\log v_\alpha - (p-2)\log t_\alpha + \log\{(v_\alpha^2 - p)\log \lambda \\
&\qquad + 4 + o(1)\} - \log\{4 + o(1)\}\Big] \qquad \text{(A.57)} \\
&= o(1)
\end{aligned}
$$

as $\alpha \to 0$, using the fact that $t_\alpha \to \infty$ as $\alpha \to 0$ and $t_\alpha \le v_\alpha$. Since $|v_\alpha - t_\alpha| \le v_\alpha - t_\alpha^2/v_\alpha$, (A.57) implies that $v_\alpha - t_\alpha \to 0$ as $\alpha \to 0$. □

FOOTNOTES

[1]I thank Inpyo Li for carrying out the Monte Carlo results reported in Section 5. I also thank Jean–Marie Dufour, Bruce Hansen, and Werner Ploberger for helpful comments. I gratefully acknowledge research support from the National Science Foundation through grant number SES–8821021.

[2]See Kim and Siegmund (1989), Chu (1989), Hansen (1989), and Banerjee, Lumsdaine, and Stock (1989) for analyses of structural change with unknown change point in linear regression models with deterministically or stochastically trending regressors. Also see Zivot and Andrews (1989) for an analysis of a unit root test against the alternative of a deterministic trend with an unknown break point. The results of the above papers are quite complementary to the results given in this paper, because they allow for trending regressors, which are not allowed here, but they only apply to linear regression models, whereas nonlinear models are considered here.

We also note that in the context of a linear regression model with non–trending regressors the results of the above papers are less general than those of the present paper. The results of Kim and Siegmund (1989) apply to a simple regression model. The results of Chu (1989) do not cover tests of partial structural change in the regression parameters. They also do not restrict the change point to be away from 0 and 1 as is necessary for the asymptotics to hold for the Wald, Lagrange multiplier, and likelihood ratio statistics. The results of Hansen (1990) are for Wald tests only. They do not include Wald tests of pure structural change in the regression parameters—the intercept is taken to be constant across the sample under the alternative. Nor do they include Wald tests of partial structural change in a subvector of the regression parameters other than the intercept. The results of Banerjee, Lumsdaine, and Stock (1989) do not apply to models without one or more stochastically trending regressors. None of the above papers covers tests of pure structural change in the regression parameters and error variance parameters, which are covered by the present paper.

[3]Although the paper concentrates on statistics of the form (1.1), the results of the paper apply more generally to statistics of the form $g(\{W_T(\pi) : \pi \in \Pi\})$ for arbitrary continuous function g (and likewise for $LM_T(\cdot)$ and $LR_T(\cdot)$). Depending upon the alternatives of interest, one may want to use a function g that differs from the "sup" function. For example, one might consider test statistics of the form $\int_\Pi h(W_T(\pi), \pi)d\lambda(\pi)$ for some function h and some measure λ.

[4]The existence of the limit uniformly over $\Theta_1 \times \Theta_3 \times \mathcal{T}_0$ means that

$$\sup_{(\theta_1,\theta_3,\tau_1)\in\Theta_1\times\Theta_3\times\mathcal{T}_0}\left|\frac{1}{T}\Sigma_1^T Em_{rt}(\theta_1, \theta_3, \tau_1) - m_r(\theta_1, \theta_3, \tau_1)\right| \to 0 .$$

[5]For example, see Billingsley (1968, p. 157) for the definition of *asymptotically independent increments*.

The two conditions stated are sufficient for a multivariate invariance principle, because (i) tightness of $\{\nu_T(\cdot)'\alpha : T \geq 1\}$ for each elementary unit v_1+v_3–vector α implies tightness of $\{\nu_T(\cdot) : T \geq 1\}$, (ii) asymptotically independent increments plus weak convergence of $\nu_T(\pi_2) - \nu_T(\pi_1)$ $\forall 0 \leq \pi_1 < \pi_2 \leq 1$ is sufficient for joint convergence of all the finite dimensional distributions of $\{\nu_T(\cdot)\}$, and (iii) weak convergence of $\nu_T(\cdot)'\alpha$ to $\nu(\cdot)'\alpha$ $\forall\alpha$ implies weak convergence of $(\nu_T(\pi_2) - \nu_T(\pi_1))'\alpha$ to $(\nu(\pi_2) - \nu(\pi_1))'\alpha$ $\forall\alpha$, $\forall 0 \leq \pi_1 < \pi_2 \leq 1$ which, in turn, implies weak convergence of $\nu_T(\pi_2) - \nu_T(\pi_1)$ to $\nu(\pi_2) - \nu(\pi_1)$ using the Cramér–Wold device (re the latter, see Billingsley (1968, p. 49)).

[6]Under the assumptions, $(\hat{M}_r'\hat{D}_r\hat{M}_r)^{-1}$ may exist only with probability $\to 1$. When $\hat{M}_r'\hat{D}_r\hat{M}_r$ is singular, a g–inverse can be used in place of the inverse. Similar comments apply elsewhere below.

[7]For the third case, note that $H(M(\pi)'D(\pi)M(\pi))^{-1}H' = HM(\pi)^{-1}D(\pi)^{-1}(M(\pi)')^{-1}H'$ $= H_*M_*(\pi)^{-1}[D(\pi)^{-1}]_*(M_*(\pi)')^{-1}H_*' = H_*(M_*(\pi)'D_*(\pi)M_*(\pi))^{-1}H_*'$ provided $D_*(\pi)^{-1} = [D(\pi)^{-1}]_*$, where the second equality uses Lemma A–3 of the Appendix.

[8]By definition, this means that η is Riemann integrable on $[0,\pi]$ $\forall\pi \in \Pi \cup \{1\}$ and $\frac{1}{T}\Sigma_1^{T\pi}\eta(t/T) \to \int_0^\pi \eta(s)ds$ uniformly over $\pi \in \Pi \cup \{1\}$ as $T \to \infty$.

[9]For the CUSUM test, there is no single critical value since the test rejects if $|CUSUM_t|$ exceeds a line with a given intercept and slope for some t. This test was size corrected by adjusting the intercept of the rejection line while holding its slope equal to the values used in Tables 3A–3C.

REFERENCES

Andrews, D. W. K. (1987): "Consistency in Nonlinear Econometric Models: A Generic Uniform Law of Large Numbers," *Econometrica*, 55, 1465–1471.

_______ (1988): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," Cowles Foundation Discussion Paper No. 877R, Yale University.

_______ (1989a): "Asymptotics for Semiparametric Econometric Models: I. Estimation," Cowles Foundation Discussion Paper No. 908, Yale University.

_______ (1989b): "Asymptotics for Semiparametric Econometric Models: II. Stochastic Equicontinuity," Cowles Foundation Discussion Paper No. 909, Yale University.

_______ (1989c): "Asymptotics for Semiparametric Econometric Models: III. Testing and Examples," Cowles Foundation Discussion Paper No. 910, Yale University.

_______ (1989d): "Generic Uniform Convergence," Cowles Foundation Discussion Paper No. 940, Yale University.

_______ (1989e): "Testing for Parameter Instability and Structural Change with Unknown Change Point in Linear and Nonlinear Econometric Models," unpublished manuscript, Cowles Foundation, Yale University.

Andrews, D. W. K. and R. C. Fair (1988): "Inference in Nonlinear Econometric Models with Structural Change," *Review of Economic Studies*, 55, 615–640.

Banerjee, A., R. L. Lumsdaine, and J. H. Stock (1989): "Recursive and Sequential Tests for a Unit Root: Theory and International Evidence," unpublished manuscript, Kennedy School of Government, Harvard University.

Billingsley, P. (1968): *Convergence of Probability Measures*. New York: Wiley.

Brown, R. L., J. Durbin, and J. M. Evans (1975): "Techniques for Testing the Constancy of Regression Relationships over Time," *Journal of the Royal Statistical Society, Series B*, 37, 149–192.

Chow, Y. S. and H. Teicher (1978): *Probability Theory: Independence, Interchangeability, Martingales*. New York: Springer–Verlag.

Chu, C.–S. J. (1989): "New Tests for Parameter Constancy in Stationary and Nonstationary Regression Models," unpublished manuscript, Department of Economics, University of California–San Diego.

Davies, R. B. (1977): "Hypothesis Testing when a Nuisance Parameter is Present Only under the Alternative," *Biometrika*, 64, 247–254.

_______ (1987): "Hypothesis Testing when a Nuisance Parameter is Present Only under the Alternative," *Biometrika*, 74, 33–43.

DeLong, D. M. (1981): "Crossing Probabilities for a Square Root Boundary by a Bessel Process," *Communications in Statistics—Theory and Methods*, A10 (21), 2197–2213.

Eberlein, E. (1986): "On Strong Invariance Principles under Dependence Assumptions," *Annals of Probability*, 14, 260–270.

Garbade, K. (1977): "Two Methods of Examining the Stability of Regression Coefficients," *Journal of the American Statistical Association*, 72, 54–63.

Hansen, B. E. (1990): "Testing for Structural Change of Unknown Form in Models with Non–stationary Regressors," unpublished manuscript, Department of Economics, University of Rochester.

Hansen, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.

Hawkins, D. L. (1987): "A Test for a Change Point in a Parametric Model Based on a Maximal Wald–type Statistic," *Sankhya*, 49, 368–376.

Hawkins, D. M. (1977): "Testing a Sequence of Observations for a Shift in Location," *Journal of the American Statistical Association*, 72, 180–186.

Hendry, D. F. (1989): *PC–GIVE: An Interactive Econometric Modelling System*, Version 6.0. Oxford: University of Oxford.

Herrndorf, N. (1984): "An Invariance Principle for Weakly Dependent Sequences of Random Variables," *Annals of Probability*, 12, 141–153.

James, B., K. L. James, and D. Siegmund (1987): "Tests for a Change–point," *Biometrika*, 74, 71–83.

Kim, H. J. and D. Siegmund (1989): "The Likelihood Ratio Test for a Change–point in Simple Linear Regression," *Biometrika*, 76, 409–423.

Kontrus, K. (1984): "Monte Carlo Simulation of Some Tests of Parameter Constancy in the Linear Model" (in German), Diplomarbeit Thesis, Department of Econometrics and Operations Research, Technical University of Vienna.

Krämer, W., W. Ploberger, and R. Alt (1988): "Testing for Structural Change in Dynamic Models," *Econometrica*, 56, 1355–1369.

Krämer, W. and H. Sonnberger (1986): *The Linear Regression Model under Test*. Heidelberg: Physica–Verlag.

Krishnaiah, P. R. and B. Q. Miao (1988): "Review about Estimation of Change Points," in *Handbook of Statistics*, Vol. 7, ed. by P. R. Krishnaiah and C. R. Rao. New York: Elsevier.

Leybourne, S. J. and B. P. M. McCabe (1989): "On the Distribution of Some Test Statistics for Coefficient Constancy," *Biometrika*, 76, 169–177.

McLeish, D. L. (1975): "Invariance Principles for Dependent Variables," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebeite*, 32, 165–178.

_______ (1977): "On the Invariance Principle for Nonstationary Mixingales," *Annals of Probability*, 5, 616–621.

Newey, W. K. (1989): "Uniform Convergence in Probability and Stochastic Equicontinuity," unpublished manuscript, Department of Economics, Princeton University.

Nyblom, J. (1989): "Testing for the Constancy of Parameters over Time," *Journal of the American Statistical Association*, 84, 223–230.

Phillips, P. C. B. and S. N. Durlauf (1986): "Multiple Time Series Regression with Integrated Processes," *Review of Economic Studies*, 53, 473–495.

Ploberger, W. and W. Krämer (1986): "The Local Power of the CUSUM and CUSUM of Squares Tests," unpublished manuscript, Technische Universität, Vienna.

Ploberger, W., W. Krämer, and K. Kontrus (1989): "A New Test for Structural Stability in the Linear Regression Model," *Journal of Econometrics*, 40, 307–318.

Pollard, D. (1984): *Convergence of Stochastic Processes.* New York: Springer–Verlag.

Pötscher, B. M. and I. R. Prucha (1989): "A Uniform Law of Large Numbers for Dependent and Heterogeneous Data Processes," *Econometrica*, 57, 675–683.

Quandt, R. E. (1958): "The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes," *Journal of the American Statistical Association*, 53, 873–880.

_______ (1960): "Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes," *Journal of the American Statistical Association*, 55, 324–330.

Roy, S. N. (1953): "On a Heuristic Method of Test Construction and Its Use in Multivariate Analysis," *Annals of Mathematical Statistics*, 24, 220–238.

Roy, S. N., R. Gnanadesikan, and J. N. Srivastava (1971): *Analysis and Design of Certain Quantitative Multiresponse Experiments.* Oxford: Pergamon Press.

Sen, P. K. (1980): "Asymptotic Theory of Some Tests for a Possible Change in the Regression Slope Occurring at an Unknown Time Point," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebeite*, 52, 203–218.

_______ (1981): *Sequential Nonparametrics: Invariance Principles and Statistical Inference.* New York: Wiley.

Talwar, P. P. (1983): "Detecting a Shift in Location," *Journal of Econometrics*, 23, 353–367.

van Zwet, W. R. and J. Oosterhoff (1967): "On the Combination of Independent Test Statistics," *Annals of Mathematical Statistics*, 38, 659–680.

Wooldridge, J. M. and H. White (1988): "Some Invariance Principles and Central Limit Theorems for Dependent Heterogeneous Processes," *Econometric Theory*, 4, 210–230.

Zacks, S. (1983): "Survey of Classical and Bayesian Approaches to the Change–point Problem: Fixed Sample and Sequential Procedures of Testing and Estimation," in *Recent Advances in Statistics* ed. by M. H. Rivzi, J. S. Rustagi, and D. Siegmund. New York: Academic Press.

Zivot, E. and D. W. K. Andrews (1989): "Further Evidence on the Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," Cowles Foundation Discussion Paper No. 944, Yale University.

TABLE 1

ASYMPTOTIC CRITICAL VALUES FOR TESTS OF PARAMETER INSTABILITY WITH $\Pi = [.15, .85]$[a]

	Significance Level			
Degrees of Freedom (p_0)	1%	2.5%	5%	10%
1	12.3	10.3	8.7	7.2
2	15.3	13.4	11.7	10.1
3	18.3	16.0	14.2	12.3
4	20.7	18.2	16.3	14.4
5	22.6	20.4	18.4	16.4
6	24.8	22.2	20.2	18.1
7	26.2	23.8	21.9	19.7
8	29.0	26.3	24.1	21.6
9	30.1	27.6	25.5	23.2
10	32.5	29.6	27.2	24.8
11	33.8	31.0	28.8	26.3
12	35.7	32.9	30.6	28.0
13	37.4	34.5	32.1	29.4
14	38.9	35.9	33.5	30.9
15	41.0	38.0	35.4	32.5
16	42.8	39.2	36.7	34.0
17	44.1	40.8	38.3	35.5
18	45.4	42.3	39.8	37.0
19	46.9	43.9	41.3	38.5
20	48.4	45.2	42.6	39.7

[a]See Comment 3 following Theorem 4 for the definition of these critical values.

TABLE 2

SIMULATED FINITE SAMPLE SIGNIFICANCE LEVELS

Sample Size	Test Statistic	1%	5%	10%
T = 30	Sup F	2.3 (.07)	7.1 (.11)	11.3 (.14)
	CUSUM	.25 (.02)	2.5 (.07)	5.8 (.10)
	Fluctuation	.01 (.00)	.53 (.03)	1.9 (.06)
T = 60	Sup F	1.3 (.05)	4.9 (.10)	8.8 (.13)
	CUSUM	.49 (.03)	3.2 (.08)	7.1 (.12)
	Fluctuation	.20 (.00)	1.6 (.03)	4.1 (.06)
T = 120	Sup F	1.0 (.04)	4.6 (.09)	8.6 (.13)
	CUSUM	.64 (.04)	.39 (.09)	8.1 (.12)
	Fluctuation	.40 (.03)	2.7 (.07)	6.2 (.11)
T = 240	Sup F	.91 (.09)	4.3 (.20)	8.4 (.28)
	CUSUM	.62 (.08)	3.8 (.19)	8.3 (.28)
	Fluctuation	.65 (.08)	3.4 (.18)	7.5 (.26)
T = 1000	Sup F	.96 (.10)	4.6 (.21)	9.2 (.29)
	CUSUM	1.1 (.10)	4.6 (.21)	9.2 (.29)
	Fluctuation	.90 (.09)	4.3 (.20)	9.3 (.29)

TABLE 3–A
SIMULATED POWER USING 5% ASYMPTOTIC CRITICAL VALUES (T = 30)

	$b\backslash\varphi =$	0°	36°	54°	90°	0°	36°	54°	90°
		(a) Sup F Test				(b) CUMSUM Test			
$\pi^* = .15$	4.8	.18	.17	.18	.17	.17	.09	.05	.01
	7.2	.36	.34	.35	.37	.35	.19	.08	.01
	9.6	.59	.60	.60	.59	.61+	.32	.11	.00
	12.0	.82	.83	.81	.82	.80	.48	.15	.00
$\pi^* = .3$	4.8	.35	.33	.32	.35	.16	.09	.04	.01
	7.2	.68	.64	.62	.67	.38	.19	.07	.01
	9.6	.90	.87	.87	.90	.62	.32	.09	.00
	12.0	.99	.99	.98	.99	.83	.46	.11	.00
$\pi^* = .5$	4.8	.43	.40	.40	.41	.09	.05	.02	.01
	7.2	.77	.76	.76	.74	.22	.11	.03	.00
	9.6	.96	.96	.96	.96	.39	.16	.04	.00
	12.0	.99	1.00	1.00	1.00	.58	.23	.05	.00
$\pi^* = .7$	4.8	.35	.35	.35	.34	.03	.02	.02	.01
	7.2	.68	.71	.68	.66	.05	.03	.01	.01
	9.6	.91	.92	.93	.91	.10	.04	.01	.00
	12.0	.99	.99	.99	.99	.14	.05	.01	.00
$\pi^* = .85$	4.8	.18	.17	.16	.17	.02	.02	.02	.02
	7.2	.37	.35	.34	.35	.02	.02	.02	.02
	9.6	.61	.59	.58	.57	.02	.01	.01	.01
	12.0	.82	.82	.81	.81	.02	.01	.01	.00
		(c) Fluctuation Test				(d) Midpoint F Test			
$\pi^* = .15$	4.8	.00	.00	.01	.01	.08	.07	.08	.07
	7.2	.01	.01	.01	.01	.09	.09	.10	.09
	9.6	.01	.01	.01	.02	.11	.11	.11	.10
	12.0	.02	.01	.01	.02	.12	.13	.13	.11
$\pi^* = .3$	4.8	.04	.02	.02	.05	.19	.18	.17	.18
	7.2	.14	.05	.05	.14	.34	.30	.31	.32
	9.6	.37	.12	.11	.37	.50	.47	.45	.48
	12.0	.67	.21	.20	.66	.67	.61	.62	.64
$\pi^* = .5$	4.8	.16	.10	.11	.16	.54+	.53+	.53+	.51+
	7.2	.53	.31	.26	.49	.87+	.89+	.87+	.85+
	9.6	.87	.59	.54	.84	.98+	.98+	.99+	.99+
	12.0	.98	.83	.81	.99	1.00+	1.00=	1.00=	1.00=
$\pi^* = .7$	4.8	.10	.07	.07	.09	.20	.20	.20	.18
	7.2	.32	.20	.19	.33	.35	.36	.37	.33
	9.6	.69	.42	.41	.65	.51	.53	.54	.50
	12.0	.91	.67	.65	.90	.67	.69	.70	.65
$\pi^* = .85$	4.8	.01	.01	.01	.01	.07	.06	.07	.07
	7.2	.02	.01	.01	.02	.09	.08	.08	.09
	9.6	.04	.01	.02	.05	.10	.09	.10	.11
	12.0	.06	.01	.02	.08	.13	.10	.11	.12

TABLE 3–B
SIMULATED POWER USING 5% ASYMPTOTIC CRITICAL VALUES (T = 60)

	$b\backslash\varphi=$	0°	36°	54°	90°	0°	36°	54°	90°
		(a) Sup F Test				(b) CUMSUM Test			
$\pi^* = .15$	4.8	.17	.15	.15	.18	.19+	.11	.07	.02
	7.2	.37	.35	.34	.36	.41+	.24	.11	.02
	9.6	.65	.60	.59	.65	.67+	.41	.18	.01
	12.0	.88	.84	.84	.87	.86	.61	.27	.01
$\pi^* = .3$	4.8	.33	.34	.32	.31	.21	.13	.07	.02
	7.2	.67	.69	.67	.69	.49	.27	.13	.01
	9.6	.93	.93	.92	.92	.75	.48	.19	.01
	12.0	1.00	.99	.99	.99	.94	.67	.27	.00
$\pi^* = .5$	4.8	.41	.41	.40	.40	.13	.07	.04	.02
	7.2	.77	.77	.80	.79	.31	.16	.06	.01
	9.6	.98	.97	.96	.96	.60	.29	.10	.01
	12.0	1.00	1.00	1.00	1.00	.80	.47	.14	.00
$\pi^* = .7$	4.8	.32	.32	.33	.34	.04	.03	.02	.02
	7.2	.67	.67	.69	.69	.10	.05	.03	.01
	9.6	.92	.92	.90	.92	.20	.09	.03	.01
	12.0	.99	.99	.99	.99	.33	.14	.03	.00
$\pi^* = .85$	4.8	.17	.19	.20	.17	.03	.03	.03	.03
	4.2	.37	.42	.41	.37	.03	.03	.02	.02
	9.6	.65	.71	.71	.66	.04	.03	.02	.01
	12.0	.88	.91	.91	.88	.05	.03	.02	.01
		(c) Fluctuation Test				(d) Midpoint F test			
$\pi^* = .15$	4.8	.02	.03	.03	.04	.08	.08	.08	.08
	7.2	.06	.04	.05	.07	.11	.10	.11	.12
	9.6	.13	.07	.08	.13	.16	.14	.15	.17
	12.0	.25	.10	.10	.24	.21	.18	.20	.24
$\pi^* = .3$	4.8	.15	.12	.12	.16	.20	.22	.22	.22
	7.2	.46	.33	.31	.46	.43	.42	.42	.41
	9.6	.81	.63	.63	.80	.63	.66	.66	.64
	12.0	.97	.87	.86	.97	.81	.83	.83	.85
$\pi^* = .5$	4.8	.32	.24	.24	.32	.56+	.58+	.56+	.56+
	7.2	.73	.59	.58	.76	.90+	.90+	.90+	.91+
	9.6	.96	.87	.88	.95	.99+	.99+	.99+	1.00+
	12.0	1.00=	.99	.98	1.00=	1.00=	1.00=	1.00=	1.00=
$\pi^* = .7$	4.8	.20	.16	.14	.20	.20	.22	.22	.22
	7.2	.54	.39	.40	.58	.42	.42	.42	.42
	9.6	.86	.68	.70	.87	.63	.65	.65	.65
	12.0	.98	.90	.90	.98	.81	.83	.83	.85
$\pi^* = .85$	4.8	.03	.04	.05	.05	.08	.10	.10	.08
	7.2	.11	.09	.10	.12	.12	.14	.14	.13
	9.6	.25	.20	.19	.26	.16	.21	.21	.19
	12.0	.49	.33	.32	.50	.23	.28	.29	.24

TABLE 3–C
SIMULATED POWER USING 5% ASYMPTOTIC CRITICAL VALUES (T = 120)

	$b\backslash\varphi =$	0°	36°	54°	90°	0°	36°	54°	90°
		(a) Sup F Test				(b) CUMSUM Test			
$\pi^* = .15$	4.8	.14	.16	.16	.17	.19+	.12	.07	.03
	7.2	.35	.35	.38	.40	.43+	.27	.14	.03
	9.6	.64	.65	.66	.68	.69+	.47	.23	.03
	12.0	.88	.88	.89	.89	.89+	.67	.34	.02
$\pi^* = .3$	4.8	.30	.30	.33	.33	.21	.12	.07	.03
	7.2	.67	.67	.68	.70	.52	.29	.13	.03
	9.6	.91	.93	.93	.93	.81	.56	.24	.02
	12.0	1.00	.99	1.00	.99	.95	.78	.36	.02
$\pi^* = .5$	4.8	.37	.37	.38	.38	.13	.08	.05	.03
	7.2	.78	.78	.77	.78	.36	.21	.08	.03
	9.6	.97	.97	.97	.97	.67	.39	.15	.02
	12.0	1.00	1.00	1.00	1.00	.88	.61	.24	.01
$\pi^* = .7$	4.8	.31	.32	.31	.30	.06	.05	.04	.03
	7.2	.67	.68	.67	.70	.12	.07	.05	.03
	9.6	.94	.93	.93	.92	.26	.13	.06	.02
	12.0	.99	.99	.99	1.00	.45	.22	.07	.02
$\pi^* = .85$	4.8	.14	.14	.15	.17	.04	.03	.03	.03
	7.2	.36	.38	.38	.39	.04	.04	.03	.03
	9.6	.66	.65	.66	.67	.06	.04	.04	.03
	12.0	.88	.87	.87	.88	.08	.05	.03	.02
		(c) Fluctuation Test				(d) Midpoint F Test			
$\pi^* = .15$	4.8	.06	.06	.06	.06	.07	.08	.08	.08
	7.2	.12	.11	.11	.14	.11	.12	.13	.12
	9.6	.27	.19	.20	.31	.18	.19	.19	.20
	12.0	.50	.35	.37	.54	.26	.27	.27	.29
$\pi^* = .3$	4.8	.22	.19	.18	.24	.19	.21	.21	.23
	7.2	.58	.46	.47	.61	.43	.43	.44	.45
	9.6	.88	.76	.78	.89	.69	.68	.67	.68
	12.0	.99	.95	.95	.99=	.87	.86	.88	.86
$\pi^* = .5$	4.8	.38+	.31	.30	.40+	.55+	.54+	.54+	.54+
	7.2	.80+	.66	.67	.80+	.90+	.90+	.91+	.91+
	9.6	.97=	.92	.92	.98+	.99+	1.00+	1.00+	1.00+
	12.0	1.00=	1.00=	1.00=	1.00=	1.00=	1.00=	1.00=	1.00=
$\pi^* = .7$	4.8	.25	.20	.19	.26	.19	.20	.20	.22
	7.2	.64	.48	.48	.65	.42	.43	.44	.45
	9.6	.91	.80	.79	.90	.69	.68	.67	.68
	12.0	.99=	.96	.95	1.00=	.87	.87	.87	.86
$\pi^* = .85$	4.8	.07	.06	.06	.08	.07	.08	.09	.08
	7.2	.16	.12	.13	.16	.12	.12	.12	.13
	9.6	.38	.24	.24	.39	.18	.18	.18	.19
	12.0	.65	.44	.42	.66	.26	.27	.28	.29

TABLE 4–A
SIMULATED POWER USING 5% SIZE ADJUSTED CRITICAL VALUES (T = 30)

	$b\backslash\varphi =$	0°	36°	54°	90°	0°	36°	54°	90°
		(a) Sup F Test				(b) CUMSUM Test			
$\pi^* = .15$	4.8	.13	.13	.14	.14	.24+	.16+	.09	.04
	7.2	.30	.28	.28	.30	.49+	.29+	.14	.02
	9.6	.53	.53	.53	.53	.72+	.46	.20	.02
	12.0	.78	.78	.76	.77	.86+	.62	.25	.01
$\pi^* = .3$	4.8	.30	.27	.26	.29	.24	.15	.08	.03
	7.2	.61	.57	.56	.60	.48	.28	.11	.01
	9.6	.88	.83	.83	.86	.72	.44	.16	.01
	12.0	.98	.98	.97	.97	.89	.58	.19	.00
$\pi^* = .5$	4.8	.35	.35	.34	.33	.14	.09	.05	.03
	7.2	.72	.70	.69	.69	.30	.16	.05	.01
	9.6	.93	.94	.94	.94	.49	.24	.08	.01
	12.0	.99	.99	.99	1.00	.69	.34	.08	.00
$\pi^* = .7$	4.8	.28	.29	.29	.26	.06	.04	.04	.03
	7.2	.62	.64	.62	.58	.09	.05	.03	.01
	9.6	.88	.88	.90	.89	.15	.07	.03	.01
	12.0	.98	.98	.98	.98	.21	.09	.02	.00
$\pi^* = .85$	4.8	.13	.12	.12	.13	.04	.05	.05	.05
	7.2	.30	.27	.27	.30	.04	.04	.04	.04
	9.6	.54	.53	.52	.52	.03	.03	.03	.03
	12.0	.77	.77	.76	.76	.02	.02	.02	.02
		(c) Fluctuation Test				(d) Midpoint F Test			
$\pi^* = .15$	4.8	.07	.07	.07	.07	.08	.07	.08	.07
	7.2	.08	.08	.08	.08	.09	.09	.10	.09
	9.6	.11	.10	.09	.11	.11	.11	.11	.10
	12.0	.13	.10	.11	.12	.12	.13	.13	.11
$\pi^* = .3$	4.8	.23	.16	.17	.21	.19	.18	.17	.18
	7.2	.51	.34	.32	.51	.34	.30	.31	.32
	9.6	.81	.55	.56	.78	.50	.47	.45	.48
	12.0	.95	.79	.75	.95	.67	.61	.62	.64
$\pi^* = .5$	4.8	.47+	.38+	.36+	.43+	.54+	.53+	.53+	.51+
	7.2	.82+	.72+	.70+	.80+	.87+	.89+	.87+	.85+
	9.6	.97+	.92	.92	.98+	.98+	.98+	.99+	.99+
	12.0	1.00+	.99=	.99=	1.00=	1.00+	1.00+	1.00+	1.00+
$\pi^* = .7$	4.8	.32+	.31+	.31+	.33+	.20	.20	.20	.18
	7.2	.71+	.61	.59	.66+	.35	.36	.37	.33
	9.6	.92+	.84	.86	.93+	.51	.53	.54	.50
	12.0	.99+	.97	.97	.99+	.67	.69	.70	.65
$\pi^* = .85$	4.8	.09	.08	.09	.10	.07	.06	.07	.07
	7.2	.15	.12	.13	.16	.09	.08	.08	.09
	9.6	.28	.16	.17	.28	.10	.09	.10	.11
	12.0	.48	.23	.22	.46	.13	.10	.11	.12

TABLE 4–B
SIMULATED POWER USING 5% SIZE ADJUSTED CRITICAL VALUES (T = 60)

	$b \backslash \varphi =$	0°	36°	54°	90°	0°	36°	54°	90°
		(a) Sup F Test				(b) CUMSUM Test			
$\pi^* = .15$	4.8	.17	.15	.16	.18	.24+	.15=	.09	.05
	7.2	.38	.35	.34	.36	.48+	.31	.15	.03
	9.6	.65	.61	.59	.65	.74+	.48	.23	.02
	12.0	.88	.84	.84	.87	.90+	.67	.33	.02
$\pi^* = .3$	4.8	.33	.34	.33	.31	.26	.16	.09	.04
	7.2	.67	.69	.67	.69	.54	.32	.15	.02
	9.6	.93	.93	.92	.93	.79	.55	.24	.02
	12.0	1.00	.99	.99	.99	.94	.72	.33	.01
$\pi^* = .5$	4.8	.41	.41	.40	.40	.17	.10	.06	.03
	7.2	.77	.77	.80	.79	.37	.20	.08	.02
	9.6	.98	.97	.96	.96	.65	.35	.12	.01
	12.0	1.00	1.00	1.00	1.00	.84	.53	.18	.01
$\pi^* = .7$	4.8	.33	.32	.33	.35	.07	.05	.04	.04
	7.2	.67	.67	.69	.69	.13	.07	.04	.02
	9.6	.92	.92	.90	.92	.24	.11	.04	.02
	12.0	.99	.99	.99	.99	.39	.18	.05	.01
$\pi^* = .85$	4.8	.17	.19	.20	.17	.05	.05	.04	.04
	7.2	.37	.42	.41	.37	.05	.04	.04	.04
	9.6	.65	.71	.71	.67	.05	.04	.03	.03
	12.0	.88	.91	.91	.88	.06	.04	.03	.02
		(c) Fluctuation Test				(d) Midpoint F Test			
$\pi^* = .15$	4.8	.08	.07	.08	.09	.08	.08	.08	.08
	7.2	.15	.12	.12	.16	.11	.10	.11	.12
	9.6	.29	.19	.18	.29	.16	.14	.15	.17
	12.0	.46	.28	.26	.47	.21	.18	.20	.24
$\pi^* = .3$	4.8	.30	.27	.25	.31=	.20	.22	.22	.22
	7.2	.66	.56	.56	.67	.43	.42	.42	.41
	9.6	.91	.84	.83	.91	.63	.66	.66	.64
	12.0	1.00=	.96	.96	.99=	.81	.83	.83	.85
$\pi^* = .5$	4.8	.48+	.43+	.42+	.49+	.56+	.58+	.56+	.56+
	7.2	.85+	.77=	.79	.87+	.90+	.90+	.90+	.91+
	9.6	.99+	.97=	.96=	.98+	.99+	.99+	.99+	1.00+
	12.0	1.00=	1.00=	1.00=	1.00=	1.00=	1.00=	1.00=	1.00=
$\pi^* = .7$	4.8	.71+	.31	.30	.38+	.20	.22	.22	.22
	7.2	.71+	.60	.63	.75+	.42	.42	.42	.42
	9.6	.94+	.86	.87	.95+	.63	.65	.65	.65
	12.0	1.00+	.98	.97	1.00+	.81	.83	.83	.85
$\pi^* = .85$	4.8	.13	.12	.12	.12	.08	.10	.10	.08
	7.2	.25	.25	.24	.25	.12	.14	.14	.13
	9.6	.46	.40	.40	.49	.16	.21	.21	.19
	12.0	.72	.60	.61	.73	.23	.28	.29	.24

TABLE 4–C
SIMULATED POWER USING 5% SIZE ADJUSTED CRITICAL VALUES (T = 120)

	$b\backslash\varphi =$	0°	36°	54°	90°	0°	36°	54°	90°
		(a) Sup F Test				(b) CUMSUM Test			
$\pi^* = .15$	4.8	.14	.17	.17	.17	.22+	.15	.09	.04
	7.2	.37	.36	.38	.42	.47+	.30	.16	.03
	9.6	.65	.67	.67	.69	.73+	.50	.26	.03
	12.0	.88	.89	.90	.90	.91+	.71	.38	.03
$\pi^* = .3$	4.8	.31	.31	.33	.35	.24	.15	.08	.04
	7.2	.68	.68	.69	.71	.55	.32	.15	.03
	9.6	.92	.93	.94	.93	.82	.60	.26	.03
	12.0	1.00	1.00	1.00	.99	.96	.79	.40	.02
$\pi^* = .5$	4.8	.38	.38	.39	.39	.15	.10	.07	.04
	7.2	.79	.79	.78	.79	.40	.22	.10	.03
	9.6	.97	.97	.97	.98	.69	.43	.17	.03
	12.0	1.00	1.00	1.00	1.00	.89	.65	.26	.02
$\pi^* = .7$	4.8	.32	.33	.32	.32	.07	.06	.05	.04
	7.2	.68	.69	.69	.71	.14	.09	.06	.03
	9.6	.94	.93	.93	.93	.29	.15	.07	.03
	12.0	1.00	.99	.99	1.00	.48	.25	.09	.02
$\pi^* = .85$	4.8	.15	.15	.16	.17	.05	.04	.04	.04
	7.2	.38	.39	.39	.40	.06	.05	.04	.04
	9.6	.68	.66	.67	.68	.07	.05	.04	.03
	12.0	.89	.88	.88	.88	.09	.06	.04	.03
		(c) Fluctuation Test				(d) Midpoint F Test			
$\pi^* = .15$	4.8	.09	.10	.10	.11	.07	.08	.08	.08
	7.2	.20	.17	.18	.22	.11	.12	.13	.12
	9.6	.39	.30	.32	.42	.18	.19	.19	.20
	12.0	.64	.48	.50	.68	.26	.27	.27	.29
$\pi^* = .3$	4.8	.32+	.27	.27	.34	.19	.21	.21	.23
	7.2	.68=	.58	.60	.72+	.43	.43	.44	.45
	9.6	.93+	.87	.87	.94+	.69	.68	.67	.68
	12.0	.99	.98	.99	1.00+	.87	.86	.88	.86
$\pi^* = .5$	4.8	.47+	.43+	.40+	.48+	.55+	.54+	.54+	.54+
	7.2	.86+	.78	.77	.87+	.90+	.90+	.91+	.91+
	9.6	.99+	.96	.96	.99+	.99+	1.00+	1.00+	1.00+
	12.0	1.00=	1.00=	1.00=	1.00=	1.00=	1.00=	1.00=	1.00=
$\pi^* = .7$	4.8	.35	.29	.28	.35	.19	.20	.20	.22
	7.2	.73+	.61	.62	.73+	.42	.43	.44	.45
	9.6	.95+	.88	.88	.95+	.69	.68	.67	.68
	12.0	1.00=	.98	.98	1.00=	.87	.87	.87	.86
$\pi^* = .85$	4.8	.11	.11	.11	.13	.07	.08	.09	.08
	7.2	.25	.21	.20	.25	.12	.12	.12	.13
	9.6	.50	.37	.37	.52	.18	.18	.18	.19
	12.0	.76	.57	.58	.77	.26	.27	.28	.29